Keele UNIVERSITY

Choice sequences and knowledge states: extending the notion of finite information to produce a clearer foundation for intuitionistic analysis

James F. Appleby

Submitted in partial fulfilment of the requirements of the degree of Doctor of Philosophy

in Mathematics

School of Computing and Mathematics, University of Keele

December, 2017

Abstract

There are currently four major formal foundational systems for intuitionistic analysis: LS, CS (both in Troelstra 1977), FIM (Kleene and Vesley 1965) and the derivable $FIRM\text{-}INT$ (Moschovakis 2016). All of these systems rely on different universes of choice sequences and different conceptions of what a choice sequence is. There is a strong common ground between these systems as they use the same very restrictive notion of finite information when dealing with these choice sequences – the notion of restricting ourselves to initial segments. This text extends the notion of a choice sequence given in Fletcher (1998) and uses it to construct a generalised system capable of expressing results about intensional properties of choice sequences. This is achieved by constructing a language capable of representing intensional first order restrictions on choice sequences (the language of knowledge states) and their relations to other sequences. This extended system allows us to formulate a notion of lawlessness that evades a series of paradoxes highlighted in Fletcher (1998), allows us to prove a generalised form of open data and offers additional clarity to other key areas of the theory. When a certain set of restrictions are applied to this extended theory (extensionality and a second order restriction on knowledge states) we obtain a system suitable for the foundation of analysis.

Acknowledgements

I could never have gotten here alone; that I made it here at all is thanks to a great many people, and would that I could mention them all. Alas I am only allowed to submit one volume, not two.

First and foremost I would like to thank Dr Peter Fletcher, my supervisor. Taking me on was a risk; a 2:1 student without a masters rarely manages to submit. No man could ask for a better teacher; Peter has proven patient, understanding and persistent. He asked the questions no-one else thought to ask, answered the questions I was afraid to ask, and pushed me to produce work of a quality I never thought myself capable of.

I've asked a great many questions during the production of this text, and while Peter has fielded a great many, I feel compelled to also thank Professor Joan Moschovakis for her eternal patience in answering my constant stream of questions about her work.

To my good friends Gavin Hedley, Christopher Murphy and Alexandra Stefanescu, you have my deepest thanks for reminding me that there are still things to enjoy in life and listening to my constant stream of unintelligible mathematics; your support kept me going during those dark times of unprovable results and faulty axioms.

To my long suffering girlfriend Ana Rocha, thank you for staying for this long, dark journey and reminding me constantly that there is light at the end of the tunnel.

For their donating their services as proofreaders, I would like to thank Simi and Tanvir; as well as 'Sam' and 'Pam' of 'PaperTrue Proofreading'.

Finally, my greatest thanks goes to Firzeen and John Appleby. I can think of no other parents who would support their son for so long while he chased his dream. I dedicate this work to you both, in thanks to you, mother, and in memory of you, father.

Contents

0. Introduction

Where does mathematics come from? The answer to this fundamental question has been debated for millennia and to date no clear answer has been given. One potential answer, is the *constructivist foundation of mathematics;* the notion that all mathematical objects are the result of mental construction, or computation. This foundational school emerged in the early 20^{th} century, and is divided into three main branches – Bishop's constructivism [fully expounded in Bishop (1967)], Russian recursive constructivism [a clear outline is given in the third chapter of Bridges and Richman (1987)], and Brouwer's Intuitionism [Troelstra and Dalen (1988) offers an excellent exposition of the conventional theory].

In the past one hundred and ten years the logical school of Intuitionism has made great leaps – from its roots in Brouwer's thesis (Brouwer 1907) and its early formalisation at the hands of Heyting (1930) to the strong foundational systems for analysis first presented in Kleene and Vesley (1965), Troelstra (1977) and Moschovakis (1987). The systems presented in these three works have been solidly refined over the years and now blaze a clear path from Brouwer's informal foundations to his strong analytic results. A central component in each of these systems is the notion of choice sequences – a notion derived from Brouwer's second act of intuitionism.

Brouwer first introduced choice sequences into his theory in Brouwer (1918) to resolve a central issue on his theory for the foundations of mathematics; specifically, he uses them to bridge the gap between the countable rationals and the uncountable reals. In Brouwer's works choice sequences are functions of type $N \mapsto N$ whose inner workings may not be entirely lawlike, i.e. not governed by an algebraically expressed function; for example, a sequence of coin tosses or rolls of a die. The challenge with working with choice sequences is that they are considered to be in a continual state of growth; we cannot treat them as completed objects. Because of this unusual property we are able to only work some finite collection of information we have about a choice sequence at a given time (so a finite sequence of coin tosses or die rolls to use our earlier example). This notion of finite information is a central one that we feel has not been considered in enough detail, and it is here that we feel that the existing theory could use some refinement.

Each of the foundational systems for analysis given above seeks to satisfy the goals raised in Dummett (1977); closure under continuous operations and a continuity axiom strong enough to obtain analysis (BC-N, given in §2.4.2). While they each meet both of these conditions the method in which each does so leaves room for improvement.

We will provide our enhancements by offering illumination on the following key questions.

- Do the current notions of lawless choice sequences need refinement and, if so, how can we go about this refinement?
- Is it possible to define a formal language for the intensional components of choice sequences and can we impose any ordering on these intensional components?
- If so, is it possible to define a foundational system that includes this formalisation of the intensional components of choice sequences?
- What concessions do we make to obtain analysis; just how severe are these concessions?

To accomplish this we do the following.

We first lay out foundational ideas in §1 and §2 and an introductory overview on the various notions of a choice sequence is given in §3.2.

Secondly, we offer a full overview of the four main systems (FIM, LS, CS and $FIRM$-INT) in §3.3 – §3.6; look closely at a group of paradoxes derived in Fletcher (1998) with respect to the treatment of choice sequences in two of these systems (LS and $FIRM$-INT) in §3.7 – §3.8 and, finally, explore the ideas exposited in a parallel branch of mathematics developed from Brouwer's ideas (the theory of the creative subject) in §3.8.

Thirdly, we devote a chapter to presenting a new notion for dealing with intensional information, the theory of knowledge states. §4.2 outlines the informal ideas; §4.3 and §4.4 outline our most basic types of knowledge state and provide formal syntax; §4.5 defines an intensional ordering for our knowledge states and deals with matters of equality outside of intensional identity; §4.6 explores some concepts useful for our foundation of analysis including our refined notion of lawlessness; §4.7 defines a species of functions capable of acting on knowledge states; building on the ideas of §4.7 and §4.2 we define a stronger continuity axiom in §4.8; in §4.9 we expose our notion of lawlessness to Fletcher's paradoxes and, finally, in §4.10 we provide a notion of bar induction in the language of knowledge states.

Fourthly, we construct our formal system in §5.1 and demonstrate how we obtain analysis from it under a certain set of clearly defined restrictions in §6.1, §6.2 and §6.3.

Fifthly, we offer a purely computational method of evaluating the functions we require for analysis in §6.4 and devote the entirety of §7.2 and §7.3 to summarising our works and suggesting further avenues for exploration.

1. Preliminaries

§1.1. Notational Conventions and Definitions

This section will quickly outline both the language and notational devices for all of the informal chapters. We will be very brief here and some concepts (such as choice sequences) will only be explained in more detail in later sections. In these instances we will provide a section reference should the reader wish to skip ahead a little to obtain more detail on a particular concept.

We will use w, x, y and z to denote natural numbers and i and j to denote natural numbers in the context of an index. We will use N to denote the natural numbers and write $x \in N$ to reinforce that x is a natural number.

We write $\equiv$ to denote identity, and $=$ for equality.

n and m will be used to denote finite sequences of natural numbers, we will use the symbol $\langle\rangle$ to denote the empty sequence. When we wish to refer to a specific element of a sequence we will write $n(i)$, where i is the desired index of the element. We write $len(n)$ to denote the length of the sequence n. Given two finite sequences $m = \langle m_1, m_2, ..., m_x\rangle$ and $n = \langle n_1, n_2, ..., n_y\rangle$ we define $m * n \equiv \langle m_1, m_2, ..., m_x, n_1, n_2, ..., n_y\rangle$. We write $n * x$ and $x * n$ to mean $n * \langle x\rangle$ and $\langle x\rangle * n$ respectively.

f and g will be used to denote constructive (the usual notion of function) functions on natural numbers, that is functions defined in some constructive way. The usual symbol, $\circ$, will be used to denote function composition.

μ and ν will be used to denote non-constructive functions of type $N \mapsto N$, better known as choice sequences. When we wish to refer to a specific element of a choice sequence we will write $\mu(i)$, where i is the desired index of the element. We write $n \subset \mu$ to denote that $\forall x \leq len(n)[n(x) = \mu(x)]$; we will sometimes refer to this as 'n being an *initial segment* of

μ'. To denote the initial segment of μ of length $x+1$ we will write $\bar{\mu}(x)$. We write M to denote the universe of choice sequences (see §3.1 for more detail).

When speaking of tuples of natural numbers, functions, finite sequences and non-constructive functions we will sometimes abbreviate $x_0, x_1, ..., x_i$ to $\underline{x}$, $f_0, f_1, ..., f_i$ to $\underline{f}$, $n_0, n_1, ..., n_i$ to $\underline{n}$ and $\mu_0, \mu_1, ..., \mu_i$ to $\underline{\mu}$. We write $|\underline{x}|$, $|\underline{f}|$, $|\underline{n}|$ and $|\underline{\mu}|$ to denote the arity of the tuple; so $|x_0, x_1, x_2| = 3$. When mentioned in a formula, unless otherwise mentioned, tuples abbreviated in this way are assumed to be of equal arity.

We we define the *flooring function* $\lfloor x \rfloor = max(y \in Z \mid m \leq x)$

p, p_1 and p_2 will denote the *Cantor pairing function* and its inverses respectively. We give these below as they will be useful to us later on to encode finite sequences as natural numbers.

$$p(x,y) = \tfrac{1}{2}(x+y)(x+y+1)+y$$

$$p_1(z) = \lfloor \tfrac{\sqrt{8z+1}-1}{2} \rfloor - z + \frac{(\lfloor \frac{\sqrt{8z+1}-1}{2} \rfloor)^2 + \lfloor \frac{\sqrt{8z+1}-1}{2} \rfloor}{2}$$

$$p_2(z) = z - \frac{(\lfloor \frac{\sqrt{8z+1}-1}{2} \rfloor)^2 + \lfloor \frac{\sqrt{8z+1}-1}{2} \rfloor}{2}$$

§1.2. Graphs and Trees

A *(directed) graph* is defined as a pair of sets of objects, **nodes** and **(directed) edges**. *Nodes* are a primitive concept and *edges* are ordered pairs of nodes. We say that an edge (n_1, n_2) connects n_1 (the *parent)* and n_2 (the *child).* If one were to draw an arrow representing an edge then the direction of the arrow would be from parent to child.

A graph is called *simple* when a node that has an edge between itself and itself does **not** exist **and**, given any pair of nodes, there will be at most one edge connecting them.

A *node labelled* graph is a pair of objects – a graph and a function mapping the nodes of the graph to natural numbers.

An *edge labelled* graph is a pair of objects – a graph and a function mapping the edges of the graph to natural numbers.

A *directed walk* is a non-empty sequence of (directed) edges $\langle (n_1, n_2), (n_2, n_3), ... \rangle$. A directed walk *connects* two nodes n_0 and n_x if it is of the form $\langle (n_0, n_1), (n_1, n_2), ..., (n_{x-1}, n_x) \rangle$.

A *directed path* is a **directed** walk where each edge occurs only once and no node (save for the first in the case of a cycle) is visited more than once.

A graph is called *connected* if given any two nodes there exists an **undirected** path between them.

A *cycle* is an **undirected** path from a node to itself.

A *(rooted directed) tree* is a directed graph where every node, save the root, has one parent and one unique node (the *root*) has no parent and has a unique **directed** path to any other node.

A *leaf* is a node with no children.

A *branch* is a directed path starting at the root of a directed rooted tree that only ends if it reaches a leaf. A branch is infinite iff it never reaches a leaf.

§1.3. Spreads and Fans

A *spread* is a rooted directed edge labelled tree with the additional restrictions that every node must have **at least** one child (a leafless tree) and, given any node, the edges to its children are uniquely labelled. Each branch of the tree forms an infinite sequence of natural numbers if one lists its edge labels in order.

A particular spread S is defined by a constructive function s called a *spread law* which, given any finite sequence (list of edge labels of a finite directed path), determines whether or not it is admissible to the spread (the initial part of a branch in the spread).

A spread law s takes the form

$$s(n) = \begin{cases} 0 \; \textit{if } n \textit{ is the initial part of a branch in the spread represented by } s \\ 1 \; \textit{otherwise} \end{cases}$$

A simple example would be the spread defined by the spread law given below.

Example 1.3.1

$$s(n) = \begin{cases} 0 \; \textit{if all elements of } n \textit{ are even} \\ 1 \; \textit{otherwise} \end{cases} \text{(the spread of even numbers).}$$

If $n = \langle 2,4,6\rangle$ then $s(n) = 0$; however, if $n = \langle 2,4,5\rangle$ then $s(n) = 1$ as it fails to meet the conditions for $s(n) = 0$. Examples of branches in the spread are $\langle 2,4,6,8,10,...\rangle$ and $\langle 16,4,162,10,16,...\rangle$.

More formally, to qualify as a spread law a constructive function s must meet the following conditions.

S1. $\forall n[s(n) = 0 \vee s(n) = 1]$

S2. $s(\langle\rangle) = 0$

S3. $\forall n[s(n) = 0 \iff \exists x[s(n * x) = 0]]$

When we say $\mu \in s$ (μ is a branch in the spread defined by spread law s) we actually mean $\forall x \; s(\bar{\mu}(x)) = 0$.

A *fan* is a spread where every node has a finite number of children.

A spread law s that satisfies the following additional condition defines a fan.

F1 $\forall n_{s(n)=0} \exists x[\forall y > x[s(n * y) = 1]]$

We pause to quickly define a spread that will be useful in the next section below.

s_{uni} $\forall n[s_{uni}(n) = 0]$ (the universal spread which every sequence is a member of)

§1.4. A Topological Definition of Continuity

This section will briefly outline the topological definitions required to define the notion of continuous operation used in this text. We are specifically interested in continuous operations mapping choice sequences to natural numbers. The reason the operations need to be continuous is that discontinuous operations can only act on completed objects, and choice sequences do not fall into this category. For a more in-depth argument as to why operations from choice sequences to natural numbers are continuous the writer recommends Van Atten and Van Dalen (2002 pp.14).

For this section we shall use some temporary notational conventions, and we give these below.

u, v - sets.

$\subset$ - the conventional subset relation.

$\cup$, $\cap$ - the conventional union and intersection.

$\rightarrow$ - 'to' as in 'maps to'.

$f[U]$ - f acting on every element of U individually.

With these temporary conventions in place we will now continue with our topological definitions.

A set of sets T is a *topology* of a set X iff

1. $\emptyset$, $X \in T$
2. $\forall u, v \in T(u \cap v \in T)$
3. Given any arbitrary collection of elements of T, their union must also be in T.

Given that T is a topology of X then (X, T) is a *topological space.* For any $u \subseteq X$, we call u *open* in (X, T) iff $u \in T$.

A set of sets B_a is a *basis* of a topology T iff every open set in T is a union of sets contained in B_a.

Given two topological spaces (X, T_X) and (Y, T_Y) and a (constructive) function $f : X \to Y$,

> We say that f is *continuous* at $x_0 \in X$, iff for each open set $v \in T_Y$ containing $f(x_0)$ there exists an open set $u \in T_X$ containing x_0 such that $f[u] \subset v$.
>
> If f is continuous at every point in X then f is a *continuous mapping* from (X, T_X) to (Y, T_Y).

Two topologies that are of interest to us are the following.

> The topological space (N, T_N) where N is the set of natural numbers and the basis for T_N is all sets of the form $\{x\}$ where $x \in N$ (i.e. all singleton sets of natural numbers).
>
> The topological space (B, T_B) where B is the universe of all choice sequences and the open sets of T_B are every set of the form $\{\mu \mid n \subset \mu\}$; n ranging over the set of all natural numbers encoding a finite sequence (all sequences sharing an initial segment).

The notion of *continuous operation* which is of interest to us is a **continuous mapping** from (B, T_B) **to** (N, T_N). We call this B*-continuity* for the duration of the text to avoid confusion with the notion of continuity we devise in §4.7. One may also define B-continuity as follows; $\psi : B \mapsto N$ is B-continuous iff $\exists x \in N \forall \mu \in B \forall \nu \in B[\bar{\mu}(x) = \bar{\nu}(x) \to \psi(\mu) = \psi(\nu)]$

§1.5. Real Numbers as Infinite Sequences

This section will briefly outline a method of constructing real numbers from infinite sequences. This text favours the method laid out in Dummett (1977 §3.2) with some slight modifications (we shall use sequences of integers rather than sequences of rationals). All results and definitions in this section are borrowed from Dummett (1977).

Given some infinite sequence of **integers** μ, we write $\langle\mu\rangle$ to denote the sequence of rationals derived from μ by applying the following operation.

$\langle\mu\rangle(x) = \frac{\mu(x)}{2^{x+1}}$

We say that μ is a *real number generator* (RNG) iff $\forall x[|\mu(x+1) - 2\mu(x)| \leq 1]$; we call the universe of all such sequences M_{RNG}.

All such sequences can be shown to Cauchy sequences as follows.

Given any μ satisfying $\forall x[|\mu(x+1) - 2\mu(x)| \leq 1]$ we have that

$$\forall x[|\mu(x+1) - 2\mu(x)| \leq 1]$$

$$\iff \forall x[|\frac{\mu(x+1)}{2^{x+2}} - \frac{2\mu(x)}{2^{x+2}}| \leq \frac{1}{2^{x+2}}]$$

$$\iff \forall x[|\frac{\mu(x+1)}{2^{x+2}} - \frac{\mu(x)}{2^{x+1}}| \leq \frac{1}{2^{x+2}}]$$

$$\iff \forall x[|\langle\mu\rangle(x+1) - \langle\mu\rangle(x)| \leq \frac{1}{2^{x+2}}] \qquad (1)$$

Given any x and any y,

$$\begin{aligned}
|\langle\mu\rangle(x+y) - \langle\mu\rangle(x)| &= |\langle\mu\rangle(x+y) - \langle\mu\rangle(x+1) + \langle\mu\rangle(x+1) - \langle\mu\rangle(x)| \\
&\leq |\langle\mu\rangle(x+y) - \langle\mu\rangle(x+1)| + |\langle\mu\rangle(x+1) - \langle\mu\rangle(x)| \\
&= |\langle\mu\rangle(x+y) - \langle\mu\rangle(x+2) + \langle\mu\rangle(x+2) - \langle\mu\rangle(x+1)| \\
&\quad + |\langle\mu\rangle(x+1) - \langle\mu\rangle(x)| \\
&\ldots \\
&\leq |\langle\mu\rangle(x+y) - \langle\mu\rangle(x+y-1)| + \ldots + |\langle\mu\rangle(x+1) - \langle\mu\rangle(x)| \\
&\leq \frac{1}{2^{x+y+2}} + \ldots + \frac{1}{2^{x+2}} \qquad (by\ 1) \\
&= \frac{2^{y-1} - 1}{2^{x+y+1}} \\
&< \frac{1}{2^{x+1}}
\end{aligned}$$

Thus, given any $\epsilon > 0$ choose x such that $\frac{1}{2^{x+1}} < \epsilon$. Then $\forall i_{i>x} \forall j_{j>x}[|\langle\mu\rangle(i) - \langle\mu\rangle(j)|] < \frac{1}{2^{x+1}} < \epsilon$.

Hence $\langle\mu\rangle$ is a Cauchy sequence.

Each element of these sequences of rationals can be mapped to an interval as follows.

$\langle\mu\rangle(x)$ encodes the interval $[\langle\mu\rangle(x) - \frac{1}{2^{x+1}}, \langle\mu\rangle(x) + \frac{1}{2^{x+1}}]$

Hence each of these sequences encodes a series of nested converging intervals.

Given two RNGs μ and ν we define the relationship $\sim$ as follows.

$\mu \sim \nu \iff \forall x[|\mu(x) - \nu(x)| \leq 2]$

We define a *real number* as the equivalence class $\{\nu \mid \mu \sim \nu\}$ for some RNG μ. For notational convenience we will use $\hat{x}$, $\hat{y}$ and $\hat{z}$to denote real numbers rather than writing out the equivalence class continuously.

We define *equality* and the *ordering* between real numbers as follows.

$\hat{x} = \hat{y} \iff \forall\mu \in M_{RNG}[\mu \in \hat{x} \iff \mu \in \hat{y}]$

$\hat{x} < \hat{y} \iff \exists\mu \in \hat{x}\exists\nu \in \hat{y}\exists x[\nu(x) - \mu(x) > 2]$

$\hat{x} \leq \hat{y} \iff \neg(\hat{y} < \hat{x})$

We define the binary *apartness* relation # as follows.

Given two real numbers $\hat{x}$ and $\hat{y}$ we have $\hat{x}\#\hat{y} \iff \exists\mu \in \hat{x}\exists\nu \in \hat{y}[\exists x[|\mu(x) - \nu(x)| > 2]]$

We define the binary operations $+$, $*$, max and min and $-$ as follows.

$\mu + \nu = \mu'$, where

$$\mu'(x) = \begin{cases} \lfloor\frac{\mu(x+2)+\nu(x+2)}{4}\rfloor & if\ x > 0 \wedge 2\mu'(x-1) \leq \frac{\mu(x+2)+\nu(x+2)}{4} \\ \lfloor\frac{\mu(x+2)+\nu(x+2)}{4}\rfloor & if\ x = 0 \wedge \frac{\mu(x+2)+\nu(x+2)}{4} > 0 \\ \lfloor\frac{\mu(x+2)+\nu(x+2)}{4} + 1\rfloor & if\ x > 0 \wedge 2\mu'(x-1) > \frac{\mu(x+2)+\nu(x+2)}{4} \\ \lfloor\frac{\mu(x+2)+\nu(x+2)}{4} + 1\rfloor & if\ x = 0 \wedge \frac{\mu(x+2)+\nu(x+2)}{4} < 0 \\ 0 & otherwise \end{cases}$$

$\mu * \nu = \mu'$, where $\mu'(x) = \frac{\mu(y)+\nu(y)}{2^y}$ rounded to the nearest integer, and where y is the smallest number satisfying $|\mu(y) - \nu(y)| \leq 2^{y-x}$. We can guarantee the existence of such a y due to all RNGs being Cauchy sequences.

$\mu - \nu = \mu + \nu'$ where $\nu'(x) = -\nu(x)$

$max(\mu, \nu) = \mu'$ where $\mu'(x) = max(\mu(x), \nu(x))$

$min(\mu, \nu) = \mu'$ where $\mu'(x) = min(\mu(x), \nu(x))$

We define the unary operation *abs* as follows.

$abs(\mu) = \mu'$ where $\mu'(x) = max(\mu(x), -\mu(x))$.

We define the unary operation *inv* on any $\mu \# 0$ as follows.

$inv(\mu) = \mu'$, where $\mu'(x) = \lfloor \frac{2^{x+y}}{\mu(y+1)} \rfloor + 1$, and where y is sufficiently large and $\lfloor x \rfloor$ is the flooring function (rounds x down to the nearest integer).

Now that we have defined our arithmetical operations we prove the following theorem that shows that our real numbers are closed under these operations.

Theorem 1.4.1

If μ and ν are RNGs then so are $\mu + \nu$, $\mu * \nu$, $\mu - \nu$, $max(\mu, \nu)$, $min(\mu, \nu)$ and $abs(\mu, \nu)$. If $\mu \# 0$ then $inv(\mu)$ is an RNG.

Proof:

We shall demonstrate the methodology via proving the case for addition, the remaining cases are left to the reader to verify in their own time.

Define R as follows $R(x) = \begin{cases} x+1 \; if \; x + \frac{1}{2} > \lfloor x+1 \rfloor \\ x \; otherwise \end{cases}$.

Given $\mu = \nu + \nu'$ then, given any x we have that:

$$\mu(x) = R(\frac{\nu(x+2) + \nu'(x+2)}{4})$$

To keep our formulae readable define $\nu(x+2) = j$ and $\nu'(x+2) = k$ and $\frac{j+k}{4} = \alpha$

$$\mu(x) = R(\frac{\nu(x+2)+\nu'(x+2)}{4}) = R(\frac{j+k}{4}) = R(\alpha)$$

Likewise,

$$\mu(x+1) = R(\frac{\nu(x+3)+\nu'(x+3)}{4}) = \begin{cases} R(\frac{2(j+k+1)}{4}) = R(2\alpha+\frac{1}{2}) & (1) \\ OR\ R(\frac{2(j+k)+1}{4}) = R(2\alpha+\frac{1}{4}) & (2) \\ OR\ R(\frac{2(j+k)}{4}) \quad = R(2\alpha) & (3) \\ OR\ R(\frac{2(j+k)-1}{4}) = R(2\alpha-\frac{1}{4}) & (4) \\ OR\ R(\frac{2(j+k-1)}{4}) = R(2\alpha-\frac{1}{2}) & (5) \end{cases}$$

So $R(2\alpha - \frac{1}{2}) \leq \mu(x+1) \leq R(2\alpha + \frac{1}{2})$

Case (1): Assume $\mu(x+1) = R(2\alpha + \frac{1}{2})$

Depending on what $j+k\ mod\ 4$ is we obtain the following cases.

$j+k\ mod\ 4 = 0$ (so $2(j+k+1)\ mod\ 4 = 2$ and we will always round down by the definition of R)

$$\mu(x) = R(\alpha) = \lfloor \alpha \rfloor$$

$$\begin{aligned} \mu(x+1) = R(2\alpha+\frac{1}{2}) &= \lfloor 2\alpha + \frac{1}{2} \rfloor \\ &= \lfloor 2\alpha \rfloor \\ &= 2\lfloor \alpha \rfloor + 1 \\ &= 2\mu(x) + 1 \end{aligned}$$

$j + k \ mod\ 4 = 1$ (so $2(j + k + 1)\ mod\ 4 = 0$)

$$\mu(x) = R(\alpha) = \lfloor \alpha \rfloor$$

$$\mu(x+1) = R(2\alpha + \frac{1}{2}) = \lfloor 2\alpha + \frac{1}{2} \rfloor$$

$$= \lfloor 2\alpha \rfloor$$

$$= 2\lfloor \alpha \rfloor + 1$$

$$= 2\mu(x) + 1$$

$j + k \ mod\ 4 = 2$ (so $2(j + k + 1)\ mod\ 4 = 2$ and we will always round down by the definition of R). There are two cases here determined by the rounding direction of $\mu(x)$:

$$\mu(x) = R(\alpha) = \lfloor \alpha \rfloor$$

$$\mu(x+1) = R(2\alpha + \frac{1}{2}) = \lfloor 2\alpha + \frac{1}{2} \rfloor$$

$$= \lfloor 2\alpha \rfloor + 1$$

$$= 2\lfloor \alpha \rfloor + 2$$

$$= 2\mu(x) + 1$$

$$\mu(x) = R(\alpha) = \lfloor \alpha \rfloor + 1$$

$$\mu(x+1) = R(2\alpha + \frac{1}{2}) = \lfloor 2\alpha + \frac{1}{2} \rfloor$$

$$= \lfloor 2\alpha \rfloor + 1$$

$$= 2\lfloor \alpha \rfloor + 1 - 1$$

$$= 2\mu(x)$$

$j + k \ mod\ 4 = 3$ (so $2(j + k + 1) \ mod\ 4 = 0$)

$$\mu(x) = R(\alpha) = \lfloor \alpha \rfloor + 1$$

$$\begin{aligned}\mu(x+1) = R(2\alpha + \frac{1}{2}) &= \lfloor 2\alpha + \frac{1}{2} \rfloor \\ &= \lfloor 2\alpha \rfloor + 1 \\ &= 2\lfloor \alpha \rfloor + 2 \\ &= 2\mu(x)\end{aligned}$$

These cases are exhaustive and, in each instance $\mu(x+1) = R(\alpha + \frac{1}{2}) \leq 2\mu(x) + 1$.

We can show, similarly, that case (5) is bounded below by $2\mu(x) - 1$ and hence $2\mu(x) - 1 \leq \mu(x+1) \leq 2\mu(x) + 1$.

Thus, no matter which form $\mu(x+1)$ takes, $|\mu(x+1) - 2\mu(x)| \leq 1$ and hence, since x was arbitrary chosen, μ is a real number generator. ♠

$<$ is a *weak ordering* iff $\forall\hat{x}\forall\hat{y}\forall\hat{z}[(\hat{x} < \hat{y} \rightarrow \hat{x} \leq \hat{y} \wedge \hat{x} \neq \hat{y}) \wedge (\hat{x} = \hat{y} \wedge \hat{y} < \hat{z} \rightarrow \hat{x} < \hat{z}) \wedge (\hat{x} < \hat{y} \wedge \hat{y} = \hat{z} \rightarrow \hat{x} < \hat{z}) \wedge (\hat{x} < \hat{y} \wedge \hat{y} < \hat{z} \rightarrow \hat{x} < \hat{z})]$

We would like to say that $<$ is a weak ordering that also satisfies $\forall\hat{x}\forall\hat{y}[\hat{x} = \hat{y} \vee \hat{x} < \hat{y} \vee \hat{y} < \hat{x}]$ ($<$ is a *total ordering*). However, this is simply not intuitionistically valid via a weak counter-example [similar to the one presented in Brouwer (1948A), though we do not use his 'creating subject' argument].

Define $\mu \in \hat{x}$ as follows.

$$\mu(x) = \begin{cases} 0 \ if \ \forall y < x[2y + 4 \ is \ the \ sum \ of \ two \ primes] \\ \frac{2^x}{(-2)^y} \ if \ y \leq x \wedge [2y + 4 \ is \ the \ least \ number \ that \ is \ not \ the \ sum \ of \ two \ primes] \end{cases}$$

In essence, the assertion of $\mu = 0$ or $\mu \neq 0$ relies on proving Goldbach's conjecture or proving its refutation; similarly for $\mu < 0$ and $0 < \mu$. Since this is an unsolved problem we have no way to assert any of these and hence we have a counter-example to $\forall\hat{x}\forall\hat{y}[\hat{x} = \hat{y} \vee \hat{x} < \hat{y} \vee \hat{y} < \hat{x}]$.

Instead, we seek to assert something slightly intuitionistically weaker (though classically equivalent).

$<$ is a weak ordering that also satisfies $\forall\hat{x}\forall\hat{y}[(\hat{x} \not< \hat{y} \wedge \hat{x} \not> \hat{y} \to \hat{x} = \hat{y}) \wedge \forall\hat{z}[\hat{x} < \hat{y} \to (\hat{x} < \hat{z} \vee \hat{z} < \hat{y})]]$ (i.e. $<$ is a *comparative order)*

A proof of this can be found in Dummett (1977 pp.49 Theorem 27).

2. Intuitionistic Ideas and Axioms

§2.1. Intuitionism

To do the one hundred and ten years of intuitionism complete justice would require a volume thrice as thick as this thesis; it is for this reason that we present a heavily abridged account focusing only on the areas with relevance to this text. Within this section we will provide references for further reading on each topic should the reader be curious to delve deeper.

§2.1.1. An Extremely Brief History of Intuitionism

Intuitionism is a philosophy on the foundations of mathematics developed by the Dutch mathematician L.E.J Brouwer as an alternative to formalism (the notion that all mathematics is some form of "symbolic game"). His original thesis was a counter to the formalism of the 1900s, however, as time went on, his well known rivalry with Hilbert's particular brand of formalism had a heavy influence on his work (see Van Stigt 1990 for a complete history of the subject). First presented in his thesis (Brouwer 1907) and revised over the years (some key papers being Brouwer (1908, 1912, 1914, 1918, 1925, 1927, 1930 and 1981)), the theory of intuitionism was never fully formalised by its creator. The work of formalisation relies primarily on the works of Heyting (1930, 1934), Kleene (Kleene and Vesley 1965), Myhill (1966, 1968), Kreisel (1958, 1967, 1968), Troelstra (1968, 1969, 1969A, 1977, 1982), Troelstra and Dalen (1988) as well as Moschovakis (1987, 1993, 1994, 2016) to provide the formal systems we see today. Excellent histories on the early development of Intuitionism can be found in Van Stigt (1990), Troelstra (1991 pp.8-20) and Van Dalen (1999).

§2.1.2. The Two Acts of Intuitionism and Choice Sequences

All of the ideas in Intuitionism are captured in two paragraphs, named (by Brouwer) 'the acts of Intuitionism'. From these two ideas Brouwer successfully constructs a coherent framework for the foundations of mathematics. The versions of both acts given in Brouwer (1981) are included below.

First Act of Intuitionism ‘*Completely* separating mathematics from mathematical language and hence from the phenomena of language described by theoretical logic, recognising that intuitionistic mathematics is an essentially languageless activity of the mind having its origin in the perception of a move of time. This perception of a move of time may be described as the falling apart of a life moment into two distinct things, one of which gives way to the other, but is retained by memory. If the twoity thus born is divested of all quality, it passes into the empty form of the common substratum of all twoities. And it is this common substratum, this empty form, which is the basic intuition of mathematics.’ [Brouwer (1981, pp.4-5)]

Second Act of Intuitionism ‘*Admitting* two ways of creating new mathematical entities: firstly in the shape of more or less freely proceeding infinite sequences of mathematical entities previously acquired (so that, for example, infinite decimal fractions having neither exact values, nor any guarantee of getting exact values are admitted); secondly in the shape of mathematical species, i.e. properties supposable for mathematical entities previously acquired, satisfying the condition that if they hold for a certain mathematical entity, they also hold for all mathematical entities which have been defined to be “equal” to it, definitions of equality having to satisfy the conditions of symmetry, reflexivity and transitivity.’ [Brouwer (1981, pp.8)]

From these two basic intuitions Brouwer was able to deduce all the tools required for mathematics. One notion in particular, that of a choice sequence, is important in this text and is derived via the second act of intuitionism. A choice sequence in Brouwer’s writing was a complex entity and his full conception will be explored in full detail in §3.2.1; for now we provide a simpler notion.

A *choice sequence* is a ‘method’ (not necessarily a lawlike one) for generating an infinite sequence of natural numbers that may contain restrictions. A *lawless choice sequence* is a choice sequence where no restrictions, save some initial segment, are put in place. A *lawlike choice sequence* is a choice sequence defined by some constructive function; please note that we **do not** equate 'constructive' with 'recursive' (recursive implies constructive, but the converse is **not** asserted); the reasons for this will be seen when we explore continuity axioms in §2.4.2.

Some simple examples of choice sequences are the following.

Writing down successive rolls of a die.

The sequence generated by the function $\lambda x.2x$ (a 'lawlike' choice sequence).

A sequence of seemingly random numbers (a 'lawless' choice sequence).

We will leave the notion of choice sequences at that for now, as their history and details will be covered in much more detail in §3.2. A remark that must be made is that choice sequences provided Brouwer with the solution to a large problem in his formulation for analysis, namely how does one bridge the gap between the countable rationals and the uncountable reals? As seen in §1.5, infinite sequences of integers (i.e. choice sequences) provide a solution to this issue.

We also pause to make note of another notion given in Brouwer's second act of intuitionism – that of a *species*. A species is a cousin to the set in that it defines a collection of objects via their possession of some form of 'property'. For example, admission to the species of even numbers relies on checking if a number is divisible by two, and admission to the species of prime numbers relies on checking if a number has any divisors save itself and one. The key difference here is that a species is not defined in terms of its contents (as a set is) but in terms of some property, which the objects in question need to satisfy. Worthy of note is that the property need not be decidable in nature.

§2.1.3. Intuitionistic Concepts

The remainder of this section will be devoted to giving a brief outline on certain important ideas that we will make use of. Each idea will contain a short list of references for readers looking to gain a more in-depth insight into these notions.

A useful notion which, according to Troelstra (1991), is implied by Brouwer in his works (Brouwer 1908, Brouwer 1924), and was further refined by Heyting for predicate logic (Heyting 1934) and Kolmogorov (1932) for propositional logic, is the Brouwer-Heyting-Kolmogorov interpretation (*BHK interpretation)* of the logical symbols. We will adopt this interpretation as it fully formalises the intuitionistic reading of logical statements.

$A \wedge B$ is read as 'we have a proof of A and a proof of B'.

$A \vee B$ is read as 'we have a proof of A or a proof of B and we know which one we have'.

$A \to B$ is read as 'we have some method of transforming a proof of A into a proof of B'.

$\forall x \in D[A(x)]$ is read as 'we have some method of converting elements of D into a proof of A for that element'.

$\exists x \in D[A(x)]$ is read as 'we have some element y in D and a proof of $A(y)$'.

$\neg A$ is shorthand for $A \to \bot$, where there is no proof of $\bot$.

The notion of what it means to 'have a proof of' is left deliberately vague.

We list below the standard logical axioms and rules of inference which are valid under this interpretations.

$\wedge$ int $\quad A \to (B \to (A \wedge B))$

$\wedge$ elim $\quad (A \wedge B) \to A$

$\wedge$ elim $\quad (A \wedge B) \to B$

$\vee$ int $A \to (A \vee B)$

$\vee$ int $B \to (A \vee B)$

$\vee$ elim $(A \vee B) \to ((A \to C) \to ((B \to C) \to C))$

$\to$ Ax1 $A \to (B \to A)$

$\to$ Ax2 $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$

T int T

F elim $\bot \to A$

$\forall$ elim $\forall s A \to A[^t_s]$ (Provided term t is substitutable for s in A)

$\exists$ int $A[^t_s] \to \exists s A$ (Provided term t is substitutable for s in A)

MP $\frac{A \quad A \to B}{B}$

$\forall$ int $\frac{A \to B}{A \to \forall x[B]}$ (Provided x does not occur free in A)

$\exists$ elim $\frac{A \to B}{\exists x[A] \to B}$ (Provided x does not occur free in B)

In Intuitionistic logic the logical axiom of *excluded middle* $(A \vee \neg A)$ is rejected, as is its parallel $\neg\neg A \to A$. For a more in-depth discussion of these ideas the author recommends Dummett (1977 pp.9-26).

An x place predicate A is considered to be iff $\forall t_0 ... \forall t_{x-1} \forall t'_0 ... \forall t'_{x-1} [t_0 = t'_0 \wedge ... \wedge t_{x-1} = t'_{x-1} \to [A(t_0, t_1, ... t_{x-1}) \iff A(t'_0, t'_1, ... t'_{x-1})]]$.

All of the predicates explored in this work are assumed to satisfy the above extensionality clause unless explicitly indicated otherwise.

We quickly differentiate two important concepts: that of *intensionality* and *extensionality.* One way to illustrate both of these ideas is with the two functions $\lambda x.6$ and $\lambda x.2 * 3$. Now, as everyone knows, $\lambda x.6 = \lambda x.2 * 3$, both functions give identical outputs; this is the idea expressed via 'extensional equality'. *Extensional* refers to the outwardly expressed behaviour of an object; some examples being the meow of a cat, and the output of a 'black box' (a machine whose workings are obscured). We also say that the operation expressed by '$\lambda x.2*3$ (multiplying 2 by 3) is different to the operation expressed by $\lambda x.6$' (outputting 6). We write this as $\lambda x.2*3 \not\equiv \lambda x.6$ which is read 'the functions $\lambda x.2*3$ is intensionally distinct from the function $\lambda x.6$'. *Intensional* essentially refers to the intrinsic definition of an object; some examples being the exact muscular process through which a cat generates its meow, and the inner workings of a black box.

§2.2. Neighbourhood Functions

A *neighbourhood function* is a constructive function **representing** a (B-) continuous operation from choice sequences to natural numbers. Neighbourhood functions act on the finite information we have about a choice sequence and maps it to a natural number, thus allowing us to circumvent the problem of the continuous operation acting on an incomplete object (since choice sequences are never considered 'completed').

In a slight break with the conventional theory, where the outputs of neighbourhood functions are represented as encoded natural numbers, we shall use a special lattice of extended natural numbers to represent the output of neighbourhood functions. This is a matter of prudence, for while natural number encoding would work just as well, we would still need to modify the existing notion when we encounter contradictory information in the extended theory. It is the author's opinion that this lattice makes reading formulae much easier as the non-output results stand out clearly in large statements. As a final note, the conventional theory makes no use of contradictory information and we only include it here to maintain notational consistency.

We define N^* to be the set $N \cup \{\Delta, \nabla\}$ and will refer to this set as the *extended natural numbers.* For our neighbourhood functions we shall use ∇ as the output for '**not enough**

information' and Δ as the output for '**contradictory information**'.

Our notion of equality (and inequality) on N^* is the standard (extensional) one, $x = x$, $\nabla = \nabla$, $\Delta = \Delta$, $\forall x[x \neq \nabla]$, $\forall x[x \neq \Delta]$ and $\Delta \neq \nabla$.

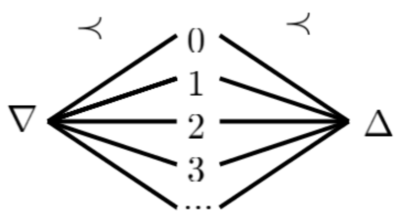
We define the binary order relation on N^*, $\prec$, below.

$\prec$-1 $\forall x \in N[\nabla \prec x \wedge x \prec \Delta]$

$\prec$-2 $\forall x^* \in N^* \forall y^* \in N^* \forall z^* \in N^*[x^* \prec y^* \wedge y^* \prec z^* \rightarrow x^* \prec z^*]$

$\prec$-3 Given any $x^* \in N^*$ and $y^* \in N^*$, $x^* \prec y^*$ iff it can be proven via $\prec$-1 and $\prec$-2.

Figure 1 below shows the resulting order.



$Fig\ 1\ :\ The\ order\ of\ N^*$

We read $A \prec B$ as 'A *precedes* B'.

We read $A \preceq B$ as 'A *precedes or is of equal position to* B', and take it to mean that $A \prec B \vee A = B$.

$(N^*, \prec)$ is easily seen to be a *lattice* as $\forall x[sup(x,x) = x \wedge inf(x,x) = x]$, $\forall x,y[x \neq y \rightarrow sup(x,y) = \Delta \wedge inf(x,y) = \nabla]$, $\forall x^* \in N^*[sup(\Delta, x^*) = \Delta \wedge inf(\Delta, x^*) = x^* \wedge sup(\nabla, x^*) = x^* \wedge inf(\nabla, x^*) = \nabla]$. So every pair of elements of N^* has a sup and an inf; hence $\prec$ is a *partial ordering* over N^*. More so since $\forall x^* \in N^*[\nabla \preceq x^* \preceq \Delta]$ $(N^*, \prec)$ is a *bounded lattice*.

The relation $\prec$ has some useful properties that will be of later use which are given below.

> $\forall x^* \in N^*[x^* \prec \Delta \iff \exists x \in N(x^* \preceq x)]$. An extended natural number precedes contradictory iff it precedes or is equal to some natural number x.

$\forall x^* \in N^*[x^* \prec x \iff x^* = \nabla]$. An extended natural number precedes some natural number x iff it is ∇.

With this lattice constructed we remind the reader that a neighbourhood function in the traditional theory is informally defined as some constructive function acting on finite sequences (representing initial segments of choice sequences) that outputs a natural number- if the finite sequence supplied is **sufficient** (long enough) to perform some continuous operation and outputs ∇ if the finite sequence is **insufficient** (too short) to perform the continuous operation.

In all conventional theories a *neighbourhood function* is a constructive (we will explain why this is so a little later) function e of type $N^\omega \mapsto (N^*\backslash\Delta)$ satisfying the following conditions.

$K_0 1$ $\forall\mu\exists x[\ e(\bar{\mu}(x)) \in N]$ (totality)

$K_0 2$ $e(\bar{\mu}(x)) \in N \rightarrow \forall y_{y>x}\ [e(\bar{\mu}(y)) = e(\bar{\mu}(x))]$ (monotonicity)

Any function satisfying these criteria is a member of the **species of neighbourhood functions** which we shall denote with K_0. We write $e \in K_0$ to abbreviate 'e is a member of K_0'.

We formally define a neighbourhood function e (of type $N^\omega \mapsto (N^*\backslash\Delta)$) as representing some continuous operation ψ (of type $B \mapsto N$)iff the following condition is met.

$$\forall\mu[\forall x[e(\bar{\mu}(x)) \preceq \psi(\mu)] \wedge \exists x[e(\bar{\mu}(x)) = \psi(\mu)]].$$

We write $\psi \sim e$ to denote an e satisfying the above for ψ.

The reader is invited to note that $e \in K_0$ alone is not sufficient to argue e lawlike. It will later be shown that, for analysis, we can prove that $e \in K_0$ does in fact imply e lawlike.

We write $e(\mu)$ to denote $e(n)$ where $n \subset \mu$ and $e(n) \in N$.

Finally, we introduce a useful notational convention adopted by the conventional theory that allows us to use neighbourhood functions to represent functions of type $M \to M$.

$$\nu = (e \mid \mu) \iff \forall x \exists y[\nu(x) = e(x * \bar{\mu}(y))]$$

§2.3. Axioms of Choice

A common concept in all existing systems is the idea of *axioms of choice,* the most common of which will be outlined here. The central idea behind the axioms of choice is to explain the meaning of the $\forall x \exists$ *something* quantifier combination, since $\exists$ has a much stronger meaning intuitionistically than it does classically. We will state the general idea behind an axiom in words before formally stating it and we will then follow this with either a brief referenced outline of the various justifications given for it, or simply a reference to a justification for it.

Our first axiom of choice is $AC\text{-}NN$ which states that 'Given that we have $\forall x \exists y A(x, y)$ then x "chooses" our y in some lawlike way'. We formally state this as follows.

$AC\text{-}NN$ $\forall x \exists y[A(x, y)] \to \exists f \forall x[A(x, f(x))]$

Troelstra (1977) claims that $AC\text{-}NN$ is 'more logical than mathematical in character' indicating that it follows directly from the meaning of $\forall x \exists y$. 'A proof of $\forall x \exists y A(x, y)$ should contain a method ('rule', 'law', 'recipe') for constructing a y [for] each x' (Troelstra, 1977 pp.5). This justification was also adopted by Dummett (1977, pp.52-53) and has endured into more modern literature.

Our second, and weakest axiom of choice would be $AC\text{-}NN!$ which states that 'Given that we have $\forall x \exists ! y A(x, y)$ then x "chooses" this **unique** y in some lawlike way'. Formally we write this as follows.

$AC\text{-}NN!$ $\forall x \exists ! y[A(x, y)] \to \exists f \forall x[A(x, f(x))]$

The justification for AC-$NN!$ is strikingly similar to that of AC-NN and is omitted here as it is rarely explicitly mentioned in the literature.

Our third axiom of choice is $AC - NF$ which states that 'Given that we have $\forall x \exists f A(x, f)$ then x "chooses" f in some lawlike way'. Formally we write this as follows.

AC-NF $\forall x \exists f[A(x, f)] \to \exists g \forall x[A(x, (g)_x)]$

where $(g)_y$ abbreviates $\forall x[(g)_y(x) = g(p(y, x))]$ (the *cross section of* g *about* x).

The justification of AC-NF is very similar to the justification of AC-NN and we will omit it here and instead refer the reader to Troelstra (1977 pp.5) and Dummett (1977, pp.52-53). For an alternative justification the reader is referred to Gielen, Swart and Veldman (1981, pp.123).

Fourth, and finally, we have the strongest axiom of choice $AC - NC$. This states that 'Given that we have $\forall x \exists \mu A(x, \mu)$ then x "chooses" μ in some lawlike way'. Formally we write this as follows.

AC-NC $\forall x \exists \mu[A(x, \mu)] \to \exists \gamma \forall x[A(x, (\gamma)_x]$

where we remind the reader that $(\gamma)_x$ abbreviates $\forall x[(\gamma)_y(x) = \gamma(p(y, x))]$.

The justification for this is, given $\forall x \exists \mu[A(x, \mu)]$ 'we first determine the value of an infinite sequence that suits [x=0], then the first value of an infinite sequence that suits [x=1]; we then continue the construction of an infinite sequence that suits [x=0] for one step and determine its second value, ... and so on'(Gielen, Swart and Veldman 1981, pp.123), essentially constructing the sequence γ using a pairing function (see §1.1).

§2.4. Continuity Axioms and Extensionality

Another unique feature found in the foundations of intuitionistic analysis are the so called *continuity axioms* and, as in the previous section, we shall expound on the most common variants here. These axioms are not consistent with classical mathematics (even combining the weakest with excluded middle derives a contradiction) and are used to explore how we evaluate predicates on choice sequences via a finite amount of information about those choice sequences. An important use of continuity axioms is the construction of strong counter-examples (two of which we shall demonstrate below); for further uses the reader is advised to consult Veldman (2001). We split this section into two parts: the first exploring a matter of great importance pertaining to the continuity schema and the second exploring the continuity schema themselves.

§2.4.1. Extensionality vs Graph Extensionality

A common theme when speaking of predicates so far has been to insist upon 'extensionality'. This notion is best illustrated with the following simple example.

Example 2.4.1.1

Given a predicate on natural numbers A, if we have $\forall x \forall y[(x = y \wedge A(x)) \rightarrow A(y)]$ then A is extensional.

All predicates on natural numbers satisfy this condition; indeed, all the predicates on rationals and reals required for analysis also satisfy it. However, as we saw in §1.5, the foundational path for analysis in intuitionism uses (equivalence classes of) choice sequences to represent real numbers. With this comes the problem of what to do with the intensional information about choice sequences; the solution chosen by the conventional theory is to 'throw this away' via insisting extensionality. More formally, it is insisted that for every predicate containing choice sequence parameters the following must hold.

Extensionality $\forall \underline{\mu} \forall \underline{\nu}[(\underline{\mu} = \underline{\nu} \wedge A(\underline{\mu})) \rightarrow A(\underline{\nu})]$

Surprisingly, this is **not** sufficient to rule out A using intensional information in some clever way that still maintains the extensionality condition. This has led to the argument that even the weakest forms of continuity that we will introduce in the next section cannot be formally justified, and that 'we can (at least at present) at most justify [the weakest continuity schema] by a plausibility argument' (Troelstra, 1977 pp.151) or 'The defence of [the weakest continuity schema] ... therefore remains on the level of a plausibility argument' (Troelstra 1983).

This is where we come to the notion of 'graph extensionality', so named in Dalen and Atten (2002). We present it formally below.

G-Extensionality A is graph extensional iff choice sequence variables 'only enter into the [predicate] A though [their] values' (Dalen and Atten 2002), i.e. choice sequences only occur in A in the form $\mu(x)$.

It is immediately clear that graph extensionality implies extensionality and that (by arguments laid out in Dalen and Atten in §3.2) the converse does not hold. If we assert that our predicates are graph extensional then there is a solid phenomenological argument for weak continuity [see Dalen and Atten (2002) and Atten (2010) for more details].

As we are presenting the conventional schema for continuity at the moment we will only insist upon the weaker notion of extensionality in our predicates, however, the reader may also read them with with the requirement of graph extensionality imposed at their leisure.

§**2.4.2. Continuity Axioms**

Our first, and weakest, continuity schema is $WC\text{-}N$ which states that 'Given we have $\forall\mu\exists x[A(\mu, x)]$ then there exists some amount of finite information (which in the conventional theory means some length of initial segment) such that any sequence ν also possessing this initial segment satisfies $A(\nu, x)$'. We present this formally below.

$WC\text{-}N$ $\forall\mu\exists x[A(\mu, x)] \rightarrow \forall\mu\exists x\exists y\forall\nu[\bar{\nu}(y) = \bar{\mu}(y) \rightarrow A(\nu, x)]$, where A satisfies extensionality and has no other free choice sequence parameter.

This notion of continuity does not require us to have defined K_0 and is derivable from all of the stronger continuity axioms explored in this section. However, as mentioned in the previous subsection, even the justification of this (the weakest possible schema for continuity) is based only on a plausibility argument. This is not to say we reject this reading of $WC\text{-}N$, but merely that we wish the reader to be aware that we are taking $WC\text{-}N$ as a (very well argued) assumption rather than an absolute derivation.

We will outline a lemma and two key results derivable from $WC\text{-}N$ below.

Lemma 2.4.2.1

$WC\text{-}N$ for disjuncts : $\forall\mu[A(\mu)\vee B(\mu)] \rightarrow \forall\mu\exists x\exists y[\forall\nu_{\bar{\nu}(y)=\bar{\mu}(y)}[A(\nu,x)]\vee\forall\nu_{\bar{\nu}(y)=\bar{\mu}(y)}[B(\nu,x)]]$

Proof: [taken from Troelstra and Dalen (1988)]

Assume $\forall\mu[A(\mu) \vee B(\mu)]$ (1)

From the intuitionistic definition of $\vee$ we have $(A\vee B)(\mu) \iff \exists x[(x = 0\wedge A(\mu))\vee (x \neq 0 \wedge B(\mu))]$

Hence (1) is equivalent to $\forall\mu\exists x[(A \vee B)(\mu, x)]$ which by $WC\text{-}N$ implies $\forall\mu\exists x\exists y\forall\nu_{\bar{\mu}(y)=\bar{\nu}(y)}[A \vee B(\nu, x)]$, which is equivalent to $\forall\mu\exists y[\forall\nu_{\bar{\nu}(y)=\bar{\mu}(y)}[A(\nu)] \vee \forall\nu_{\bar{\nu}(y)=\bar{\mu}(y)}[B(\nu)]]$

Thus discharging our hypothesis we obtain $\forall\mu[A(\mu)\vee B(\mu)] \rightarrow \forall\mu\exists x\exists y[\forall\nu_{\bar{\nu}(y)=\bar{\mu}(y)}[A(\nu,x)] \vee\forall\nu_{\bar{\nu}(y)=\bar{\mu}(y)}[B(\nu,x)]]$ as required. ♠

Theorem 2.4.2.1

$WC\text{-}N$ refutes excluded middle.

Proof:[taken from Troelstra and Dalen (1988)]

Assume excluded middle in the form $\forall\mu[\forall x[\mu(x) = 0] \vee \nexists x[\mu(x) = 0]]$.

Then by our disjunctive WC-N above we have

$$\forall \mu \exists x \exists y [\forall \nu_{\bar{\nu}(y)=\bar{\mu}(y)} [\forall x [\nu(x) = 0]] \vee \forall \nu_{\bar{\nu}(y)=\bar{\mu}(y)} [\not\forall x [\nu(x) = 0]]] \ (1)$$

Choose $\mu \equiv \lambda x.0$; then for some y with $n = \bar{\mu}(y)$, (1) implies $\forall \nu [\bar{\nu}(y) = n \rightarrow \forall x [\nu(x) = 0] \vee \forall \nu [\bar{\nu}(y) = n \rightarrow \not\forall x [\nu(x) = 0]]]$.

The first disjunct is false since no matter how large our y is we could have $n * 1 \subset \nu$.

The second disjunct is also false since $\nu \equiv \mu$ leads to an easily derived contradiction.

Since we have reached a contradiction we must conclude that our hypothesis is false, as we made no other assumptions, and thus WC-N refutes excluded middle as required. ♠

This result is a key factor in the development of intuitionistic analysis as it shows us that excluded middle must be omitted no matter how weak our continuity axiom is.

Theorem 2.4.2.2

WC-N is incompatible with $\forall \mu \exists x \forall y \exists z [T(x, y, z) \wedge U(z)]$ (*Church's Thesis*), where T is Kleene's T predicate and U is the corresponding U function.

Proof:

We offer only an informal outline of a proof here; the details may be verified by the reader at their leisure. Applying WC-N to Church's Thesis leads to obtaining an algorithm (comprised of a finite number of symbols) capable of generating every possible choice sequence starting with a certain initial segment which is, of course, impossible. ♠

This result indicates why Church's Thesis is usually rejected in intuitionistic mathematics; it denies us one of the weakest forms of continuity. There is a weaker version of continuity, WC-$N!$, which is compatible with [Church's Thesis]; in fact it can be derived from a different version of [Church's Thesis] and Markov's Principle (MP). However, it is not strong enough to produce the results we require for intuitionistic analysis (Troelstra and Dalen, 1988 pp.211).

This is why, in §2.1.2, we state we 'do not equate "constructive" with "recursive" ', to do so would be to deny ourselves even the weakest form of **useful** continuity.

Our second continuity axiom is the most widely used; it is sufficient to obtain all major results in analysis and is unproblematic unlike the strongest continuity axiom. This axiom essentially states that if we have $\forall\mu\exists x[A(\mu, x)]$ then x can be derived in some constructive (but not necessarily lawlike) way from μ. We present this formally below.

$BC\text{-}N$ $\forall\mu\exists x[A(\mu, x)] \rightarrow \exists e \in K_0\forall\mu[A(\mu, e(\mu))]$, where A satisfies extensionality and has no other free choice sequence parameters.

In more modern conventional literature, such as Troelstra and Dalen (1988), this is referred to as $C\text{-}N$ rather than $BC\text{-}N$, with $BC\text{-}N$ being reserved for $e \in K$ (K being a subset of K_0 which we will define in §2.6) rather than $e \in K_0$. We will not follow this convention in this text and instead use the traditional notation presented in Draglin (1988).

Theorem 2.4.2.3

$BC\text{-}N \vdash WC\text{-}N$

Proof:

Assume $BC\text{-}N$ and the hypothesis of $WC\text{-}N$.

Then by $BC\text{-}N$ we have $\exists e \in K_0\forall\mu[A(\mu, e(\mu))]$ where $e(\mu) = x$.

Since $e \in K_0$,

By $K_0 1$ $\forall\mu\exists y[e(\bar{\mu}(y) \in N]$, i.e. $\forall\mu\exists y\exists x[e(\bar{\mu}(y)) = x]$

By $K_0 2$, for our y above, $\forall z > y[e(\bar{\mu}(y)) = e(\bar{\mu}(z))]$. So if $\bar{\mu}(y) = \bar{\nu}(y)$ then $e(\nu) = e(\mu) = x$.

Hence we have proven that $\forall\mu\exists y\exists x\forall\nu[\bar{\mu}(y) = \bar{\nu}(y) \rightarrow A(\nu, x)]$.

Thus by discharging our assumption we obtain $\forall\mu\exists x[A(\mu,x)] \rightarrow \forall\mu\exists x\exists y\forall\nu[\bar{\nu}(y) = \bar{\mu}(y) \rightarrow A(\nu,x)]$ as required. ♠

Theorem 2.4.2.4

$WC\text{-}N \wedge \{\forall\mu\exists x[A(\bar{\mu}(x))] \rightarrow \exists e \in K_0\forall\mu\exists n \subset \mu[A(\bar{\mu}(e(n)))]\} \vdash BC\text{-}N$

Proof: [taken from Troelstra and Dalen (1988)]

Assume WC-N and $\forall\mu\exists x[A(\bar{\mu}(x))] \rightarrow \exists e \in K_0\forall\mu\exists n \subset \mu[A(\bar{\mu}(e(n)))]$ (*) and the hypothesis of BC-N.

Applying WC-N to the hypothesis of BC-N gives

$\forall\mu\exists y\exists x\forall\nu[\bar{\mu}(y) = \bar{\nu}(y) \rightarrow A(\nu,x)]$ (1)

Choose $\nu = \mu$ and define $z = p(x,y)$.

Thus (1) implies that $\forall\mu\exists z[\bar{\mu}(u_2(z)) = \bar{\nu}(u_2(z)) \rightarrow A(\mu,u_1(z))]$ (2)

Define $A'(\bar{\mu}(z)) \iff \bar{\mu}(u_2(z)) = \bar{\nu}(u_2(z)) \rightarrow A(\mu,u_1(z))$.

(2) is now equivalent to $\forall\mu\exists z[A'(\bar{\mu}(z))]$ and hence, by (*), this implies $\exists e' \in K_0\forall\mu\exists n \subset \mu[A(\bar{\mu}(e'(n)))] \iff \exists e' \in K_0\forall\mu\exists n \subset \mu[A(\mu,u_1(e'(n))]$ which (by defining $e(\mu) = u_1(e'(\mu)))$ is equivalent to $\exists e\forall\mu\exists n \subset \mu[A(\mu,e(n)]$. (3)

$\forall\mu\exists x[e'(\bar{\mu}(x) \in N)]$ by $K_0 1$ and u_1 is clearly total so if $e(\bar{\mu}(x)) \in N$ then $u_1(e(\bar{\mu}(x))) \in N$; therefore $\forall\mu\exists x[e(\bar{\mu}(x)) \in N]$.

Given any n and m such that $n \leq m$ then $e'(n) \preceq e'(m)$ by $K_0 2$. Appealing to the extensionality of our pairing operation this means that $e(n) \leq e(m)$ and hence, discharging our assumptions and quantifiers over n and m we have $\forall m\forall n[n \leq m \rightarrow e(n) \preceq e(m)]$.

Hence $e \in K_0$ and we can amend (*) to $\exists e \in K_0\forall\mu\exists n \subset \mu[A(\mu,e(n)]$.

Discharging our hypothesis of $BC\text{-}N$ yields $\forall\mu\exists x[A(\mu,x)] \to \exists e \in K_0 \forall\mu\exists n \subset \mu[A(\mu, e(n))]$ as required. ♠

This pair of theorems show how the various axioms of continuity are inter-related. It is this tangled web of inter-relations that grants this field its complexity. Changing one thing necessitates that the resulting consequences must be widely checked!

Theorem 2.4.2.5

$BC\text{-}N \vdash AC\text{-}NN$

Proof: [taken from Troelstra and Dalen (1988)]

Assume $BC\text{-}N$ and the hypothesis of $AC\text{-}NN$.

$\forall x\exists y[A(x,y)] \iff \forall\mu\exists y[A(\mu(0),y)]$ and hence by applying $BC\text{-}N$ to the RHS of this we obtain $\exists e \in K_0\forall\mu[A(\mu(0), e(\mu))]$ (1)

Choosing μ as $\lambda z.x$ allows us to rewrite (1) as $\exists e \in K_0\forall x[A(x, e(\lambda z.x))]$ (2)

Defining $f \equiv \lambda x.e(\lambda z.x)$ allows us to rewrite (2) as $\exists f\forall x[A(x, f(x))]$.

Discharging out hypothesis of $AC\text{-}NN$ yields $\forall x\exists y[A(x,y)] \to \exists f\forall x[A(x,f(x))]$ as required. ♠

This final theorem shows just how far the inter-relations spread, linking continuity axioms with choice axioms.

Our final continuity axiom is the strongest and is often considered controversial due to its disagreements with *Kripke's Schema* ($\exists\mu[\forall x[\mu(x) = 0] \equiv \neg A]$ being a weak form of this), with Dummett (1977 pp.450) even going so far as to say 'there is no intuitive reason to suppose [$BC\text{-}C$]'. This axiom essentially states that if we have $\forall\mu\exists\nu[A(\mu,\nu)]$ then ν can be derived in some constructive (but not necessarily lawlike) way from μ. We present this formally below.

BC-C $\forall \mu \exists \nu [A(\mu, \nu)] \rightarrow \exists e \in K_0 \forall \mu [A(\mu, e \mid \mu)]$, where A satisfies extensionality and has no other free choice sequence parameters.

There are many propositions about *BC-C*'s relationship with the other continuity axioms; however, there are no concrete proofs. We state the most obvious proposition below.

Proposition

$BC\text{-}C \vdash BC\text{-}N$

Reasoning:

Assume the hypothesis of *BC-N*, convert it into the form $\forall \mu \exists \nu [A(\mu, \nu(0)]$; then all we need to do is apply *BC-C* to obtain $\exists e \in K_0 \forall \mu [A(\mu, e \mid \mu)]$ and from here we just project the first element of our resulting sequence to get $\exists e \in K_0 \forall \mu [A(\mu, (e \mid \mu)(0))]$. $(e \mid \mu)(0)$ can be shown to be a K_0 function and hence obtain our desired result.

Finally,we demonstrate below just how *BC-C* conflicts with even a weak version of Kripke's Schema given in Draglin (1988 pp.135–136).

Theorem 2.4.2.6

$BC\text{-}C \wedge \exists \mu [\forall x [\mu(x) = 0] \equiv \neg A]$ (KS^-) leads to a contradiction.

Proof: [taken from Draglin (1988)]

By KS^- $\forall \mu \exists \nu [\forall x [\nu(x) = 0] \equiv \not\forall x [\mu(x) = 0]]$.

Hence by*BC-C* $\exists e \in K_0 [\forall x [(e \mid \mu)(x) = 0] \equiv \not\forall x [\mu(x) = 0]]$.

Define $\mu_0 = \lambda x.0$ and $\nu_0 = e \mid \mu_0$ and assume $\not\forall x [\nu_0(x) = 0]$.

Suppose that for some $\exists z [\nu_0(z) = S(x) > 0]$.

Then $\exists y [e(z * \bar{\mu}_0(y)) = S(S(x))]$.

Define mu_1 such that $\bar{\mu}_1(y) = \bar{\mu}_0(y)$ and $\mu_1(\nu) = 1$.

Since $\bar{\mu}_0(y) = \bar{\mu}_1(y)$ we have $(e \mid \mu_0)(z) = (e \mid \mu_1)(z) = s(x)$.

By the definition of e, because we have $\exists z[(e \mid \mu_1)(z) \neq 0]$, we also must have $\forall x[\mu_1(x) = 0]$ which contradicts the fact that $\mu_1(y) = 1$.

Thus the hypothesis is false, and hence $\forall z[\nu_0(z) = 0]$, but this contradicts our earlier hypothesis on ν_0! ♠

§2.5. Bar Induction, the Fan Theorem and the Uniform Continuity Theorem

This section is split into two parts; §2.5.1 briefly outlines three key notions of bar induction (BI_D, BI_T and BI_M) as well as the relationships between them, and §2.5.2 makes use of BI_D to prove two major theorems in intuitionistic analysis (FAN and UC).

§2.5.1. Bar Induction

A very important idea in the development of intuitionistic analysis is some axiom representing the idea of *bar induction.* From this crucial idea we derive the fan theorem and, from this, the uniform continuity theorem. This subsection will be split into two parts; the first exploring the various axioms of bar induction, the second stating and proving the fan and uniform continuity theorems.

Given some spread s and some predicate R, if $\forall \mu \in s \exists x[R(\bar{\mu}(x))]$ holds then R *bars the spread.* We say that a predicate A *is upward hereditary* to mean that $\forall n[\forall x[A(n * x)] \to A(n)]$. We are interested in three forms of bar induction – Decidable, Monotonic and Thin – but before we look closely at the types of bar induction it will be helpful to explore the general idea of bar induction.

An example of applying (decidable) bar induction is given below.

Example 2.5.1.1

Let us say we have two people playing chess and we record their moves as a sequence (every move can be encoded as natural numbers); once checkmate is reached a special move (the empty move) occurs forever after. Thus each such game of chess forms a choice sequence in the fan of chess games. A common rule used in tournament chess is the '50 move rule' (also known as the FIDE 50 rule); if 50 moves occur without a capture by either side the game is declared as a stalemate. We define R such that $R(n) \iff$ "Checkmate has been reached or stalemate has occurred".

(i) R clearly bars the spread of chess games (every game of chess eventually ends or reaches a stalemate).

(ii) R is decidable (we can always spot if a game has ended or if a stalemate has been reached). R is a property on finite sequences on moves.

Define A such that $A(n)$ iff any game with initial moves n terminates after finitely many moves.

(iii) $R(n) \to A(n)$ since if R holds, then the game has terminated in finitely many moves.

(iv) If there is a sequence of moves n such that, no matter what move you make next, the game will eventually terminate ($\forall x[A(n * x)]$ holds) then you can claim that any game beginning with the sequence of moves n also leads to eventual game termination in finitely many moves, hence $A(n)$ (A is upward hereditary).

All forms of bar induction allow us to formulate results about predicates on infinite sequences via predicates on decidable ones. Since A and R satisfy (i)-(iv), (decidable) bar induction allows us to assert $A(\langle\rangle)$, which is equivalent to the statement that "every game of chess ends in finitely many moves". This is different from the statement $R(\langle\rangle)$ which would be claiming "the game with no moves has ended", a clearly false statement.

Note, this does not assert that there is an upper bound on the finite number of moves! This example will be revisited at a later point (§2.5.2) and will be referenced there when it is.

Below are the three possible schemas of bar induction (all considered to be equivalent when taken together with $BC\text{-}N$).

Decidable bar induction (BI_D**)** requires that $\forall n[R(n) \vee \neg R(n)]$. Formally we write it as

$$
\begin{aligned}
&[\forall \mu \exists x[R(\bar{\mu}(x))] \\
&\wedge \forall n[R(n) \vee \neg R(n)] \\
&\wedge \forall n[R(n) \to A(n)] \\
&\wedge \forall n[\forall x(A(n * x)) \to A(n)]] \\
&\quad \to A(\langle\rangle)
\end{aligned}
$$

Monotonic bar induction (BI_M**)** requires that $\forall n[R(n) \to \forall x(R(n * x))]$. Formally we write it as

$$
\begin{aligned}
&[\forall \mu \exists x[R(\bar{\mu}(x))] \\
&\wedge \forall n[R(n) \to \forall x(R(n * x))] \\
&\wedge \forall n[R(n) \to A(n)] \\
&\wedge \forall n[\forall x(A(n * x)) \to A(n)]] \\
&\quad \to A(\langle\rangle)
\end{aligned}
$$

Thin bar induction (BI_T) requires that $\forall\mu\exists!x[R(\bar{\mu}(x))]$. Formally we write it as

$$
\begin{aligned}
&[\forall\mu\exists!x[R(\bar{\mu}(x))] \\
&\wedge \forall n[R(n) \rightarrow A(n)] \\
&\wedge \forall n[\forall x(A(n * x)) \rightarrow A(n)]] \\
&\quad \rightarrow A(\langle\rangle)
\end{aligned}
$$

In some systems relations exist between these notions of bar induction; these will be given as theorems below and proven.

Theorem 2.5.1.1

$BI_M \vdash BI_D$

Proof: [taken from Dummett (1977)]

Assume BI_M and assume that R and A satisfy the premise of BI_D.

Define R' and A' as follows.

$R'(n) \iff \exists m \subseteq n[R(m)]$

$A'(n) \iff A(n) \vee R'(n)$

R' is monotonic by definition.

We have $\forall\mu\exists x[R(\bar{\mu}(x))]$ by our assumptions on R, and so we have $\forall\mu\exists x[R'(\bar{\mu}(x))]$ by the definition of R'.

$R'(n) \rightarrow A'(n)$ by the definition of A'.

All that remains is to show that A' is upward hereditary and that $A'(n) \rightarrow A(n)$.

Assume that $\forall x[A'(n * x)]$. R is decidable by our hypothesis and hence R' is also decidable. Thus we only need to argue the following cases; $R'(n)$ and $\neg R'(n)$.

1. If $R'(n)$, then $A'(n)$ by the definition of $A'(n)$.

2. If $\neg R'(n)$, then $R'(n * x) \to R(n * x)$ since $\neg R(n)$.

We thus have that $R'(n * x) \to A(n * x)$ by our hypothesis on R and A.

This means we have $\forall x[A'(n * x) \to A(n * x)]$ by the definition of A'.

Hence we have $\forall x[A(n * x)] \to A(n)$ by the upward hereditaryness of A and thus $A'(n)$.

Both our cases are satisfied and hence we have $\forall x[A'(n * x)] \to A'(n)$.

R' and A' now satisfy the hypothesis of BI_M and hence we obtain $A'(\langle\rangle)$ by BI_M.

By the definition of A' we have $A'(\langle\rangle) \iff A(\langle\rangle) \vee R'(\langle\rangle)$.

If $A(\langle\rangle)$ then our work is already done for us.

If $R'(\langle\rangle)$ then $R(\langle\rangle)$ by the definition of R', and hence $A(\langle\rangle)$.

Thus $A(\langle\rangle)$ will hold regardless of which disjunct is true.

Discharging our assumptions on A and R yields BI_D as required. ♠

Theorem 2.5.1.2

$BI_D \wedge BC\text{-}N \vdash BI_M$

Proof: [taken from Dummett (1977)]

Assume BI_D and that R and A satisfy the hypothesis of BI_M.

$\forall \mu \exists x[R(\bar{\mu}(x))] \to \exists e \in K_0 \forall \mu[R(\bar{\mu}(e(\mu)))]$ by $BC\text{-}N$

We define R' as follows.

$$R'(n) \iff \exists m \subseteq n[e(m) \succ \nabla \ \wedge \ \forall m' \subset m[e(m') = \nabla] \ \wedge \ lth(n) \geq e(m)]$$

R' is decidable by definition.

Since $e \in K_0$ we have by $K_0 1$ that $\forall \mu \exists x[e(\bar{\mu}(x)) \in N]$, which is equivalent to $\forall \mu \exists x[R'(\bar{\mu}(x))]$.

Since R is monotonic by hypothesis $R'(n) \to R(n)$ because $e(n) \in N$ and hence $R'(n) \to A(n)$ by hypothesis.

Finally, we have that A is upward hereditary by hypothesis.

Hence R' and A satisfy the hypothesis of BI_D and hence we have $A(\langle\rangle)$ by BI_D.

Hence by discharging our assumptions on A and R we obtain BI_M as required. ♠

Theorem 2.5.1.3

$BI_T \vdash BI_D$

Proof:

Assume A and R meet the hypothesis of BI_D.

Define R' as follows.

$$R'(n) \iff [R(n) \wedge \neg \exists m < n[R(m)]]$$

Clearly $\forall \mu \exists ! x[R'(\bar{\mu}(x))]$ and $\forall n[R'(n) \to A(n)]$.

Hence, by BI_T, $A(\langle\rangle)$.

Discharging our assumptions on A and R yields BI_D as required. ♠

Theorem 2.5.1.4

$BI_D \vdash BI_T$

Proof:

Assume A and R satisfy the hypothesis for BI_T.

We note that $\forall\mu\exists!x[R(\bar{\mu}(x))] \to \forall\mu\exists x[R(\bar{\mu}(x))]$.

The $\exists!x$ quantifier in $\forall\mu\exists!x[R(\bar{\mu}(x))]$ implies that we can, both, construct such an x and recognise such an x as being our unique x. The only way this would be possible is if R were decidable; otherwise how would we determine if the x we had constructed is the correct one?

Thus all thin bars must be decidable and hence $\forall\mu\exists!x[R(\bar{\mu}(x))] \to \forall n[R(n) \vee \neg R(n)]$

Hence we have shown that R and A satisfy the hypothesis for BI_D and thus $A(\langle\rangle)$ by BI_D.

Discharging our hypothesis we obtain BI_T as desired. ♠

§2.5.2. The Fan and Uniform Continuity Theorems

We introduce here the proof of two major theorems in intuitionistic analysis – the *fan theorem* (FAN) and the uniform continuity theorem (UC). The proof of the latter will make use of an additional theorem, the extended fan theorem ($EFAN$). The fan theorem is derivable from all three of our schema's of bar induction; however, we shall only include a proof of it via BI_D as this is the path we shall take for our own analysis later in this text.

Theorem 2.5.2.1 : FAN

$$(fan(s)$$

$$\wedge\forall\mu \in s\exists x[\bar{\mu}(x) \in R]$$

$$\wedge\forall n[n \in \vee n \notin R])$$

$$\rightarrow\exists x\forall\mu \in s\exists y \leq x[\bar{\mu}(y) \in R]$$

Proof: [taken from Dummett (1977)]

Assume s and R satisfy the hypothesis of FAN.

Define the predicate Q and the species A as follows.

$$Q(n, x) \iff \forall\mu \in s[n \subseteq \mu \rightarrow \exists y \leq x[R(\bar{\mu}(y))]]$$

$$A(n) \iff \exists x[Q(n, x)]$$

Assume that for some given n, where $s(n) = 0$, that $\forall x_{s(n*x)=0}[A(n * x)]$ which is equivalent to $\forall x_{s(n*x)=0}\exists y[Q(n * x, y)]$ holds.

Since $fan(s)$ holds by hypothesis $\exists y'$ such that $\forall x > y'[s(n * x) = 1]$.

By AC-NN this implies that $\forall x_{s(n*x)=0}[Q(n*x, f(x))]$ where f is a constructive function.

By the definition of Q this implies that $Q(n, Max(f(x) \mid s(n * x) = 0))$ holds, which is equivalent to $\forall x[A(n * x)] \rightarrow A(n)$.

By definition $R(n) \rightarrow A(n)$ and hence A and R satisfy the hypothesis of BI_D, therefore $A(\langle\rangle)$.

This means that $\exists x\forall\mu[\langle\rangle \subset \mu \rightarrow \exists y \leq x[R(\bar{\mu}(y))]]$ which is equivalent to $\exists x\forall\mu\exists y \leq x[R(\bar{\mu}(y))]$.

Discharging our assumptions on R and s we obtain FAN as required. ♠

Before moving on to the proof of the extended fan theorem, we provide an illustration of how FAN may be applied to a real world problem below.

Example 2.5.2.1

We earlier proved that every game of chess ends in finitely many moves (see Example 2.5.1.1), however we were unable to assert that there is an upper bound on the number of chess moves (so we did not rule out games going on for arbitrarily many moves). We will now do so by making use of the fan theorem.

(i) During every (legal) game of chess there are only finitely many possible moves at a given time; thus the spread of all chess games must be a fan.

(ii) Our earlier defined R ($R(n)$ iff the game of chess represented by n has terminated either via checkmate or stalemate) satisfies both $\forall \mu \in s \exists x[\bar{\mu}(x) \in R]$ (where μ is the choice sequence representing a game of chess) and $\forall n[n \in \vee n \notin R]$).

Hence, by the fan theorem, $\exists x \forall \mu \in s \exists y \leq x[\bar{\mu}(y) \in R]$, i.e. "there is an upper bound on the number of moves a chess game may have under the FIDE 50 rule".

As a point of interest, finding this upper bound is merely an exercise in logic; there are 30 take-able pieces on the board, once the last one has been taken stalemate is declared. If we make 49 moves before taking each piece (the move being taken on the 50th move) we find that the longest possible game of chess would be $50 * 30 = 1500$ moves, i.e. our upper bound (x) is 1500. A very long afternoon of very badly played chess indeed!

Theorem 2.5.2.2 : $EFAN$

$(fan(s) \wedge \forall\mu \in s\exists x[A(\mu, x)]) \rightarrow \exists y\forall\mu \in s\exists x\forall\nu \in s_{\bar{\mu}(y)\subset\nu}[A(\nu, x)]$ where A has no other choice sequences parameters.

Proof: [taken from Dummett (1977)]

Assume the hypothesis of $EFAN$, then by $BC\text{-}N$ $\exists e \in K_0\forall\mu \in s\exists n \subset \mu[A(\mu, e(n)]$.

Define the species R as follows.

$$n \in R \iff e(n) \in N \wedge \forall m < n[e(n) = \nabla]$$

R is decidable and $\forall\mu \in s\exists x[\bar{\mu}(x) \in R]$.

Hence by FAN $\exists y\forall\mu \in s\exists x \leq y[\bar{\mu}(x) \in R]$, which is equivalent to $\exists y\forall\mu \in s\exists x \leq y[e(\bar{\mu}(x)) \in N \wedge \forall m < \bar{\mu}(x)[e(m) = \nabla]]$.

This is equivalent to $\exists y\forall\mu \in s\exists x \leq y\forall z[e(\bar{\mu}(z)) = \nabla \wedge A(\mu, e(\bar{\mu}(x)))]$.

Given any $\nu \in s$ such that $\bar{\mu}(y) \subset \nu$ then $e(\bar{\nu}(y)) = e(\bar{\mu}(y))$ and hence $A(\nu, e(\bar{\mu}(y)))$.

Thus $\exists x\forall\mu \in s\exists y\forall\nu \in s_{\bar{\mu}(y)\subset\nu}[A(\nu, x)]$.

Discharging our assumptions on s and A we obtain $EFAN$ as required. ♠

Lemma 2.5.2.1

Given a $\mu \in M_{RNG}$ and $n \subset \mu$ where $len(n) = l$.

Define the spread s as follows.

$$s(m) = 0 \iff \exists\mu \in M_{RNG}[m \subset \mu] \wedge (n \leq m \vee m \leq n)$$

Then if $\hat{x} \in [\frac{\mu(l)-1}{2^{l+1}}, \frac{\mu(l)+1}{2^{l+1}}]$ we have $\exists\mu' \in s[\mu' \in \hat{x}]$.

Proof:

See Dummett (1977 pp.117 Theorem 3) for an easily generalisable proof of this. ♠

We pause to offer a particular definition of *uniform continuous* (in case it has slipped the reader's mind).

A function f is *uniform continuous* on the interval [x,y] iff

$\forall i \exists z \forall \hat{x} \in [x,y] \forall \hat{y} \in [x,y][|\hat{x} - \hat{y}| < \frac{1}{2^z} \rightarrow |f(\hat{x}) - f(\hat{y})| < \frac{1}{2^i}]$.

Theorem 2.5.2.3 : *UC*

Any function $f : [x,y] \mapsto R$ is uniform continuous.

Proof: [taken from Dummett (1977)]

Given any $\hat{x} \in [x,y]$ then $\forall \mu \in \hat{x}[f(\langle\mu\rangle) = f(\hat{x})]$ (we remind the reader that $\langle\mu\rangle$ is defined in §1.5 and that $\hat{x}$ and $\hat{y}$ denote real numbers).

Given any i we have that, for any $\hat{x} \in [x,y]$, we can approximate $f(\hat{x})$ to within $\frac{1}{2^{i+1}}$. Formally we can write this as $\forall i \forall \mu \in M_R[|f(\langle\mu\rangle) - \frac{\mu(i)}{2^{i+1}}| < \frac{1}{2^{i+1}}]$.

The relationship between μ and $\mu(i)$ is purely extensional and hence by $EFAN$ we have that $\forall i \exists k \forall \mu \in M_R \forall \nu \in M_{R\bar{\mu}(k)\in\nu}[|f(\langle\nu\rangle) - \frac{\mu(i)}{2^{i+1}}| < \frac{1}{2^{i+1}}]$.

Suppose, for our k we have $\hat{x} \in [x,y]$ and $\hat{y} \in [x,y]$ such that $|\hat{x} - \hat{y}| < \frac{1}{2^{k+1}}$. By lemma 2.5.2.1 we have $\forall \hat{x} \exists \mu \in M_R[\mu \in \hat{x}]$, and we can construct some $\nu \in \hat{y}$ such that $\bar{\mu}(k) = \bar{\nu}(k)$. Hence,

$$|f(\hat{x}) - f(\hat{y})|$$

$$=|f(\langle\mu\rangle) - f(\langle\nu\rangle)|$$

$$\ngtr |f(\langle\mu\rangle) - \frac{\mu(i)}{2^{i+1}}| + |f(\langle\nu\rangle) - \frac{\nu(i)}{2^{i+1}}|$$

$$< \frac{2}{2^{i+1}}$$

$$= \frac{1}{2^i}$$

For any i we can find such a k so f is uniformly continuous as required. ♠

We pause here to make a crucial point; provided that we produce a notion of real number similar to the one represented in §1.5, and we can obtain the schema $BC\text{-}N$, $AC\text{-}NN$ and BI_D, we have all we require to deduce UC. This idea will play a central role in our conclusion in §6.3 and thus we provide it here for when the reader will invariably look back.

§2.6. Inductively Defined Neighbourhood Functions

In some conventional theories (IDB, LS and CS) a special species of neighbourhood functions is used in place of bar induction. This species of (lawlike) *inductively defined neighbourhood functions* is denoted by K and is defined by the following axioms.

K1 $\forall x[\lambda n.x \in K]$

The constant functions are in K.

K2 $(e(\langle\rangle) = \nabla \wedge \forall x[\lambda n.e(x * n) \in K]) \rightarrow e \in K$

If the empty sequence is insufficient to evaluate the function and every 'cross section' of the function is in K then it is also in K.

K3 $\forall e[(e \equiv \lambda n.x \vee (e(\langle\rangle) = \nabla \wedge \forall x[\lambda n.e(x * n) \in Q])) \rightarrow e \in Q] \rightarrow \forall e[e \in K \rightarrow e \in Q]]$

K is the minimal set of inductively defined neighbourhood functions.

In the conventional theory it is provable that $K \subseteq K_0$; however bar induction is obtained from the (intuitionistically unproven) hypothesis (sometimes called K-induction) that $K = K_0$. We pause to note that a parallel notion (that of a 'stump') is defined in Veldman (2006) to prove his 'improved bar theorem' (a result strong enough to derive FAN and hence UC). This is done without the need to resort to $K = K_0$, though it does rely on $AC\text{-}NC$. ♠

Theorem 2.6.1

$K \subseteq K_0$

Proof: [Taken from Troelstra and Dalen (1988)]

We set $Q = K_0$ in $K3$.

It can easily be shown that, if e is of the form $\lambda n.x$, then $e \in K_0$ (the constant functions).

Assume that for some e that $e(\langle\rangle) = \nabla \wedge \forall x[\lambda n.e(x * n) \in K_0]$.

Given any function f define $f' \equiv \lambda x.f(x+1)$, so $f \equiv f(0) * f'$ so $e(f(x+1)) = e(f(0) * \bar{f}'(x))$.

By our assumption $\lambda n.e(f(0) * n) \in K_0$ $(x = f(0))$, hence $\exists x[e(f(0) * \bar{f}'(x)) \in N]$ satisfying $K_0 1$.

For $e(n) \in N$ we must have $n = x * n'$ for some x and n'. Hence by our hypothesis $e(x * n') \in K_0$.

Thus $\forall m[e(x * n') = e(x * n' * m) = e(n * m)]$ satisfying $K_0 2$ and hence $e \in K_0$.

Hence by $K3$ $K \subseteq K_0$ as required. ♠

There is a classically valid proof that $K_0 \subseteq K$. However, we will omit this to preserve the intuitionistic flavour of our text. If the reader desires to verify such a result the author recommends referring to Troelstra and Dalen (1988 pp.227).

Our next major result will be crucial for analysis; however, we state the following lemma from Troelstra and Dalen (1988 pp. 230) beforehand, without proof, to ease our path.

Lemma 2.6.1 : IUS "Induction over Unsecured Sequences"

$$e \in K \to (\forall n[e(n) \in N \to A(n)] \wedge \forall n[\forall x[A(n * x)] \to A(n)]) \to A(\langle\rangle)$$

With this lemma, and the assumption that $K = K_0$, we may now prove the following theorem.

Theorem 2.6.2

$$K = K_0 \vdash BI_D$$

Proof: [taken from Troelstra and Dalen (1988)]

Let R and A fulfil the hypothesis of BI_D.

Define γ as follows.

$$\gamma(n) = \begin{cases} 1 \ m \leq n[R(m)] \\ \nabla \ otherwise \end{cases}$$

Clearly $\gamma \in K_0$.

Define $A'(n) \iff A(n) \vee \exists m < n[R(m)]$.

Assume that for some arbitrary n we have $\gamma(n) \in N$, then $\exists m \leq n[R(m)]$.

If $m = n$ then $R(n) \to A(n) \to A'(n)$ by definition.

If $m < n$ then $A'(n)$ by definition.

Thus $\forall n[\gamma(n) \in N \to A'(n)]$.

Assume, given some arbitrary n that $\forall y[A'(n * y)]$ holds.

This is equivalent to $\forall y[A(n * y) \vee \exists m < n * y[R(m)]]$ which can be rewritten as $\forall y[A(n * y) \vee \exists m \leq n[R(m)]]$.

This is decidable and, when considered case by case, we obtain $\forall y[A'(n * y)] \to A'(n)$.

Hence by applying our lemma IUS to A' we obtain $A'(\langle\rangle)$ which, by definition, implies $A(\langle\rangle)$.

Discharging our assumptions on A and R yields BI_D as required. ♠

It is here and now that we are in the position to argue for the 'lawlikeness' of K_0 functions, if we assume $K = K_0$. In Van Der Hoeven (1982 pp.19) K is introduced as 'the subset of the set of lawlike elements in [Baire Space] inductively defined by [$K1$-$K3$]' which clearly defines K as entirely lawlike and, hence, by the assertion of $K = K_0$, K_0 is also lawlike.

Theorem 2.6.3

$$\forall\psi[\exists e \in K[e \sim \psi] \iff \forall\mu\forall\nu[\mu = \nu \to \psi(\mu) = \psi(\nu)]]$$

Proof:

Right to left is trivial and relies on $K = K_0$ and the fact that ψ is Baire continuous. (1)

Given any ψ assume we have some $e \in K[e \sim \psi]$, and given any μ and ν assume $\mu = \nu$.

$\mu = \nu \to \forall x[\mu(x) = \nu(x)]$ and since $K \subseteq K_0$ we can assert by K_02 that $\mu = \nu \to e(\mu) = e(\nu)$ and hence since $e \sim \psi$ we have $\psi(\mu) = \psi(\nu)$.

Discharging our quantifiers over μ and ν, and assumptions about μ and ν, this gives us $\forall\mu\forall\nu[\mu = \nu \to \psi(\mu) = \psi(\nu)]$.

So $\forall x[e(\bar{\mu}(x)) = e(\bar{\nu}(x))]$ since K functions are monotonic and only ever work with initial segments.

Hence by the definition of $\sim$ we obtain $\psi(\mu) = \psi(\nu)$.

By discharging our assumption and quantification over e and quantification over ψ we obtain $\forall\psi[\exists e \in K[e \sim \psi] \to \forall\mu\forall\nu[\mu = \nu \to \psi(\mu) = \psi(\nu)]$. (2)

Hence by (1) and (2) we can assert $\forall\psi[\exists e \in K[e \sim \psi] \iff \forall\mu\forall\nu[\mu = \nu \to \psi(\mu) = \psi(\nu)]]$ as required. ♠

Before continuing on to the next section we state the following useful lemma without proof.

Lemma 2.6.2

$$\forall\psi[\psi\ B\text{-continuous} \iff \exists e \in K_0[e \sim \psi]]$$

Proof : Follows from the definition of B-continuous.♠

3. Existing Formal Systems

§3.1. Chapter Outline

This chapter explores the existing literature on systems of choice sequences – those of Brouwer (1918, 1925, 1927, 1930 and 1981), Kleene and Vesley (1965), Kreisel (1968) and Troelstra (1968, 1969, 1969A, 1977, 1982), Moschovakis (1987, 1993, 1994 2016) and the group working on the creative subject. §3.2 explores the notion of choice sequences in more detail; §3.3 constructs the formal system $PrAn$ which is used in the construction of all later systems; §3.4 constructs the formal system FIM, the system of Kleene and Vesley; §3.5 constructs the formal systems IDB, LS and CS attributed to Kreisel and Troelstra; §3.6 constructs the formal systems $FIRM\text{-}INT$ and $FIRM$ of Moschovakis; §3.7 compares the system LS with $FIRM\text{-}INT$; §3.8 briefly explores the work done on the theory of creative subject, and §3.9 offers a summary of our findings in the literature.

§3.2. Choice Sequences

This section aims to explore the various notions of choice sequences in existence today; §3.2.1 explores the notion of a choice sequence arrived at by Brouwer, §3.2.2 explores the notion of lawless and proto-lawless sequences favoured in the systems of Kreisel and Troelstra, §3.2.3 explores the notion of choice sequences introduced in Van Atten and Van Dalen (2002), §3.2.4 explores the notion of choice sequences as presented in Fletcher (1998) and §3.2.5 explores the notion of GC (*Generated Continuously*) sequences informally.

§3.2.1. Brouwer's Choice Sequences

The first part of this section draws heavily on the extremely useful accounts presented in both Troelstra (1977 appendix A1) and Troelstra (1982).

In the original form of intuitionism, first presented in Brouwer's 1907 thesis, choice sequences were absent. According to Troelstra (1977 pp.128), and all texts translated into English from that period, Brouwer (1912) first mentions choice sequences in two paragraphs (both only two paragraphs apart from each other).

> 'Let us consider the concept: 'real number between 0 and 1'. For the formalist this concept is equivalent to "elementary series of digits after the decimal point", for the intuitionist it means "law for the construction of an elementary series of digits after the decimal point, built up by means of a finite number of operations". And when the formalist creates the "set of all real numbers between 0 and 1", these words are without meaning for the intuitionist, even whether one thinks of the real numbers of the formalist, determined by elementary series of freely selected digits or of the real numbers of the intuitionist, determined by finite laws of construction.' (Brouwer 1912).

> 'If we restate the question in this form: "Is it impossible to construct infinite sets of real numbers between 0 and 1, whose power is less than that of the continuum, but greater than aleph-null?", then the answer must be in the affirmative; for the intuitionist can only construct denumerable sets of mathematical objects, and if on the basis of the intuition of the linear continuum, he admits elementary series of free selections as elements of construction, then each non-denumerable set constructed by means of it contains a subset of the power of the continuum' (Brouwer 1912).

While we agree with Troelstra's interpretation of the first paragraph; 'admitting choice sequences (= free choices of digits) seems to be the formalist's business, not a concern of the intuitionist, who deals only with lawlike sequences'; we feel compelled to add something to Troelstra's interpretation of the second quotation; 'the second quotation however shows that the intuitionist at least conceivably might use choice sequences' (Troelstra, 1977 pp.129). An equally likely alternative interpretation of the second quotation is that Brouwer rejects the notion of choice sequences as an intuitionistic concept entirely at this point. The reason for this interpretation is because Brouwer claims 'the intuitionist can only construct denumerable sets of mathematical objects' and his final point, seemingly offering hope for the concept of choice sequences, relies on 'each non-denumerable set constructed' which he claims is beyond the reach of an intuitionist.

It is only in a book review (Brouwer 1914) that Brouwer first revealed a definite interest in such objects: 'I recall that for the intuitionist only such infinite sets exist which are composed from ... and a second part, which is based on ... while each of its elements is determined from a sequence of choices of elements from a finite or countable infinite set in such a way that ...' (Brouwer 1914; translation from Troelstra 1977 pp.129).

Brouwer's first exposition involving choice sequences was written in 1917 (though the paper was only published in 1918).

> 'A *spread* [Menge] is a *law* [Gesetz], on the basis that, whenever an arbitrary number-symbol complex of the sequence ζ is chosen each of these choices either generates a particular symbol or nothing or effects the suspension [Hemmung] of the process and the definite annihilation of its results, whereby for every n after the unsuspended sequence of $n-1$ choices at least one number-symbol complex can be denoted which, if it is chosen as the n-th number-symbol complex, does *not* lead to the suspension of the process. Every sequence of symbols generated in this fashion by the spread (which hence is in general not presentable as completed) is called an *element of the spread*. For brevity we will also denote the collective type of genesis [Entstehungsart] of the elements of a spread M as *the spread M*.
>
> If different choice sequences [Wahlfolgen] always lead to different symbol sequences then the spread is called *individualised*.' (Brouwer 1918; original translation by J.F.Appleby and J.Rittburg (Unpublished)).

This is the first time Brouwer openly uses the phrase 'choice sequences' in the literature and, as can be seen, he constructs them alongside his notion of spreads. The introduction of spreads does indicate the acceptance of some form of restriction, although perhaps only an initial one. According to Troelstra (1982 pp.472), a letter from Brouwer to Heyting in 1924 is the first time a restriction on choice sequence is mentioned; interestingly, it is a second order restriction (a restriction on restrictions).

> 'Let R be the species of rationals, V the species of 'free' members of the continuum (i.e. those numbers such that in their approximation by a convergent sequence of λ-intervals the choice of the next λ-interval within the previous one remains **completely free**' (Taken from Troelstra (1982) pp.472, underlining replaced by bold).

The first formal mention of restrictions was made in Brouwer (1925) in a footnote for the definition of spreads.

> 'Including the feature of their freedom of continuation, which after each choice can be limited arbitrarily (possibly to being fully determined, but in any case according to a spread law)' (Brouwer 1925A, translation taken from Troelstra (1982) pp.473).

By this point it is safe to assume that Brouwer's conception of a choice sequence was that of a sequence of pairs – a 'chosen' element and a spread restriction (possibly empty, possibly a law), though no notion of second order restrictions was included in the original version according to Troelstra (1988). Brouwer later [in 1929 according to Heyting's note in Brouwer (1975 pp.590 note 3)] refined his footnote in the 1925A paper to allow for higher order restrictions.

> 'The **arbitrariness** of this possible restriction, which does not violate the possibility of continuation, adds a new element of arbitrariness to this choice sequence and its continuations. It is also possible to join a well-ordered species of restrictions to the spread (e.g. a restriction of the existing freedom of adding restrictions on future choices).' (Brouwer, 1925A; translation from Troelstra 1982 pp.473, underlining replaced by bold).

However he later came to doubt the idea of second order restrictions.

> 'In some former publications of the author restrictions of freedom of future restrictions of freedom, restrictions of freedom of future restrictions of freedom of future restrictions of freedom, and so on, were also admitted. But at present the author is inclined to think this admission superfluous and perhaps leading to needless complications.' (Brouwer, 1981 pp.13)

> 'However, this admission is not justified by close introspection, and moreover would endanger the simplicity and rigour of further developments.' (Brouwer, 1952B footnote pp.142)

Thus Brouwer's final conception of choice sequences was, as above, sequences of pairs of initial segments and sets of restrictions introduced at each step of the construction of the sequence. At any time during the generation of additional elements (lengthening of the initial segment) new restrictions may be added (the set of restrictions is non-static), though once present restrictions are permanent.

Brouwer's conception of a choice sequence is not the only one in use today; it is to the exploration of these other universes of choice sequences that the rest of this section lends itself.

§3.2.2. Lawless and Proto-Lawless Choice Sequences

Not everyone agreed with Brouwer's rejection of second order restrictions, or more precisely, one specific second order restriction; the injunction against further restrictions. It is Kreisel who first coined the term 'lawless sequence' in Kreisel (1968) to describe these types of sequence; previously they were named 'free sequences' or 'free arrows'.

> 'The sequences to be considered in detail in the present paper are those where the simplest kind of restriction on restrictions is made, namely some finite initial segment

> of values is prescribed, and, beyond this, no restriction is made. I expressed this idea by absolutely free [in Kreisel 1958], but shall call these sequences lawless here' (Kreisel 1968, footnote omitted).

In Troelstra (1977 pp.12) this definition was modified slightly to omit the requirement of some pre-specified initial segment.

> 'We think of a lawless sequence (of natural numbers), as a process (not a law!) of assigning values to the arguments 0,1,2 ...; at any stage, only finitely many values (i.e. an initial segment) of the sequence are known; at no stage we impose restrictions on the future possibilities for assigning values to arguments (except, in this case, the general a priori restriction that all values will be natural numbers)'

Though, in Troelstra (1983) it was refined to re-include it.

> 'initially one may specify that [our lawless choice sequence] starts with an initial segment [n]'

This re-inclusion came about by the recognition of a subset of the lawless sequences, namely the 'proto-lawless sequences'. The distinction between a lawless and proto-lawless sequence is simple; a lawless sequence has some pre-defined initial segment before any elements are generated by the process whereas a proto-lawless sequence does not. An example illustrating this distinction is given below.

Example 3.2.2.1

> If we had a die and were to cast it and write down the values we would have a proto-lawless choice sequence. If we deliberately placed the die a few times before we cast it and wrote down those deliberate placings before any casts we would have a lawless choice sequence.

A key consequence of this difference between the lawless and proto-lawless sequences is the validity of density; given any possible initial segment n, is there (proto-) lawless choice sequence beginning with that initial segment? Density is only valid for lawless choice sequences by virtue of the pre-defined initial segment and thus proto-lawless sequences cannot be said to satisfy density via this method. Further explorations as to whether proto-lawless sequences do or do not satisfy density seems to have been abandoned without a satisfactory resolution.

The use to which lawless sequences are put is largely in constructing universes of sequences suitable for analysis [Van Der Hoeven (1982) gives an excellent account of this], specifically, we use them to construct GC sequences (see §3.2.5). The reason that lawless sequences are not used directly to construct analysis is that they are not closed under continuous operations!

§3.2.3. Choice Sequences with Temporary Restrictions

A tacit assumption that many make when reading Brouwer is that any restriction imposed cannot be lifted. A paper by Van Atten and Van Dalen (2002) explored the idea of dividing restrictions [Note 195 in Van Atten (2010 pp.157) admits that the notion of allowing liftable restrictions was first mentioned in Van Dalen (1968), though not fully explored] into the so called 'provisional restrictions' defined as follows.

> '... provisional restrictions: which have the form "for an unspecified number of stages, [this restriction] holds" ' (Van Dalen (2002)).

The definitive restrictions defined are as follows.

> '... definitive restrictions: which have the from "from now on, [this restriction] holds, and it won't be revised anymore" ' (Van Dalen (2002)).

The notion of provisional restriction is an interesting one, since one feels inclined to ask 'why bother with them at all, surely they tell us nothing useful?'. Van Atten (2010 pp.108) gives answer to this by stating that 'If I tell you that I won't begin by lifting the provisional restriction [during the generation of this element], then you do know something about my next choice of a term, namely, that it has to respect this restriction. Had I imposed no restrictions, then you would not have known this'. In essence, this implies that before we generate an element we must make a statement of our intentions with regard to restrictions in place, i.e. when generating a choice sequence we must make a statement about restrictions, then choose an element, then make another statement about restrictions, then choose another element, etc.

This notion of a choice sequence does not clash with that of Brouwer and it does offer a much more 'natural' idea as to what a choice sequence is as well as providing a justification to WC-N (Van Atten, 2010 chapter 7) with the aid of phenomenological ideas.

§3.2.4. The 'Functional Interpretation' of Choice Sequences

A key issue with the theory of choice sequences is its requirement of some form of temporal notion; it is to the removal of this that the functional interpretation of Fletcher (1998) (also known as the 'black box' interpretation) lends itself.

Fletcher (1998 pp.135-136) proposes that we 'replace the picture of a [lawless] choice sequence as a process generating values one after another with the picture of a *black box,* that is, a device that generates a value when supplied with a natural number argument'. The immediate advantage of this notion is that the temporal component is entirely stripped from the theory; though it is not without its drawbacks. Fletcher (1998 pp.136) points out that 'a disadvantage of taking black boxes as the basic notion is that we need to make the assumption that the behaviour of the box is deterministic and history-independent; that is, it will always produce the same output for a certain input. This may seem like a slightly artificial assumption given that we are pretending not to know anything about what is going on inside the box'.

Given that the notion of lawless sequences presented by Kreisel and Troelstra requires us to accept a second order restriction on choice sequences, it seems no great leap to extend this notion to the production of 'black boxes' (the property that all black boxes will be temporally invariant); one could hardly call one artificial without strongly implying the same about the other since neither assumption seems more outlandish.

We pause here to note an additional point absent from Fletcher's account, namely that we can now no longer be certain if a relationship between two given choice sequences is present or not; this is due to the complete absence of intensional information (save identity) when regarding choice sequences. This proves no great problem; however, it does imply that we treat all choice sequences as 'lawless' in a stronger sense then that of Kreisel and Troelstra expounded in §3.2.2.

As a final closing remark Fletcher's notion of black box alone is not strong enough to obtain a complete theory of intuitionistic analysis, as neither continuity nor choice principles are defined in terms of black boxes, however Fletcher does write 'this is not because I reject them but because ... it is not yet clear how to incorporate them without damaging the integrity of the choice sequence or black box idea' (Fletcher, 1998 pp.139). Before this notion of choice sequences becomes a viable alternative it requires taking a little further; a notion for both continuity and choice needs defining before we can fully accept this idea of choice sequences.

§3.2.5. *GC* Choice Sequences

With the advent of Kreisel and Troelstra's system CS (covered in §3.5.3) there came a need for a notion of a choice sequence satisfying it, and it is from here that the notion of GC (Generated Continuously) sequences was first born.

The notion of GC *sequences* was first presented in Troelstra (1969 and 1969A), though an earlier less general notion only allowing unary relations was proposed in Troelstra (1968).

The notion was subjected to further analysis in both Troelstra (1977) and Dummett (1977) with Dummett offering some refinements on the notion, though these refinements did lead to an increasingly complex idea. A full outline of the modern notion of GC sequences is provided in Van Der Hoeven (1982 chapter 2); we shall only attempt to present a very informal idea here, lest we be forced to include a companion volume to this text!

The notion of GC sequences is that of sequences generated continuously, that is a sequence defined in 'stages'. At each stage sequences can have: additional elements generated; additional sequences introduced, and additional dependences (relationships) between sequences introduced. Stage by stage, we define our sequence(s) ad infinitum, adding restrictions and/or elements at each step, which is very close to the notion proposed by Brouwer.

The key difference between Brouwer's notion of a choice sequence and GC sequences is the nature of the restrictions; for GC sequences only restrictions of the form $\mu = \psi(\mu_1)$ (where ψ is some continuous operator) or $\mu = p(\mu_1, \mu_2)$ (where p is our pairing function) are allowed. Also, once a sequence is restricted in terms of other sequences, its element generation is done solely through the continuous operation; we may no longer choose elements for that sequence independently and may only freely choose the elements of sequences that are not defined in terms of any other.

Our main critique for this notion is its complexity; as Goodman (1979) remarks on Dummett's refinements; 'Dummett finishes with a notion built up from lawless sequences and continuous functional in a manner so complex that it is difficult to believe that he intends this to be the fundamental notion on which analysis to be founded'.

§3.3. The System $PrAn$

The system $PrAn$ (Primitive Recursive Analysis) forms a starting point for both classical and intuitionistic mathematics. We have adapted this system from the one given the same name in Draglin (1979 pp.127) as a common root for our formal systems. This section will construct $PrAn$ for later use. $PrAn$ treats choice sequences as a special kind of function of type $N \mapsto N$ and thus allows us to make use of the language of functions to work with these objects; this was the conventional approach at the time and it still remains valid to this day.

The language of $PrAn$ contains variables for natural numbers, denoted by the symbols x, y, z, and variables for functions, denoted by the symbols μ and ν. To make defining special functions easier we shall use Church's λ notation and the function composition operator $\circ$. The operations of intuitionistic predicate logic (the equality connective $=$, the order on the extended natural numbers $\prec$, the usual logical connectives $\wedge$, $\vee$, $\rightarrow$, $\exists$, $\forall$ and the logical constants T and $\perp$) are used for formula formation of $PrAn$. Finally our constants are 0, ∇, Δ and a finite collection of symbols to represent the atomic primitive recursive functions (f_1, f_2,...).

The following primitive recursive functions will be used frequently and are thus assigned their own abbreviations.

p will be taken to abbreviate our primitive recursive pairing function (see §1.1)

$(\mu)_x$ will be taken to abbreviate the primitive recursive *cross sections function* $\forall x[(\mu)_y = \mu(p(y, x))]$ which extracts a subdomain of a function based on a natural number input.

$S(x)$ will be taken to abbreviate the primitive recursive *successor function* $\lambda x.x + 1$.

$COMP$ will be taken to abbreviate the primitive recursive function defined as follows.

$$COMP(f, g_0, g_1, ...g_y)(x_0, ..., x_z) = f(g_0(x_0, ..., x_z), g_1(x_0, ..., x_z), ..., g_y(x_0, ..., x_z))$$

REC will be taken to abbreviate the primitive recursive function defined as follows.

$$REC(f, g)(0, x_0, ...x_y) = f(x_0, ..., x_y)$$

$$REC(f,g)(S(z),x_0,...x_y) = g(z,x_0,...,x_y,REC(f,g)((z,x_0,...x_y)))$$

We will use the characters n and m to denote any natural number term representing a finite sequence encoded by repeated applications of the pairing function p; this is done to remove the need for finite sequences in the language. For notational ease, given $n = \langle y_0, y_2, ..., y_x, .., y_z \rangle$, we shall allow the use of $n(x)$ to abbreviate applying the unpairing function p_1 on n $x+1$ times.

There are three types of terms in $PrAn$ – *numerical*, *extended numerical* and *functional*. The term formation is as follows.

$PrAn$-$T1$ The constant 0 or any natural number variable x are **numerical** terms.

$PrAn$-$T2$ Any function variable μ or any function variable representing a primitive recursive function is a **functional** term.

$PrAn$-$T3$ For any function term T and the set of i extended numerical terms $t_1, t_2, t_3, ..., t_i$, $\ulcorner T(t_1, t_2, t_3, ..., t_i) \urcorner$ is a **numerical** term.

$PrAn$-$T4$ Any function defined by Church's λ (of the form $\lambda x.\mu$ where μ is a function variable) is a **functional** term.

$PrAn$-$T5$ Given two function terms t and t', $\ulcorner t \circ t' \urcorner$ and $\ulcorner t \circ t' \urcorner$ are **functional** terms.

$PrAn$-$T6a$ Any numerical term is also an **extended numerical** term.

$PrAn$-$T6b$ Δ and ∇ are **extended numerical** terms.

Atomic formulae of $PrAn$ have the form $\ulcorner t = s \urcorner$, where t and s are both numerical, extended numerical or functional terms of the **same type** in the language of $PrAn$. The formulae of $PrAn$ are constructed by applying logical connectives and quantifiers to atomic formulae.

Given that A, B, C represent formulae in $PrAn$, we adopt the following standard notational conventions.

$\neg A$ is taken to mean $A \to F$.

A term t *occurs freely* in a formula A **iff** at least one occurrence of t in A is not bound by any quantifier.

A term t is *substitutable for a variable* k in A **iff** one of the following is true.

A is atomic.

A is $\neg B$ and t is substitutable for k in B.

A is $B \vee C$ (or $B \wedge C$ or $B \to C$) and t is substitutable for k in B and in C.

A is $\exists k_0 B$ (or $\forall k_0 B$) and **either** t does not occur freely in B **or** k_0 does not occur in t and t is substitutable for k in B.

We will write 'substitute t for s in A' as $A[{t \atop s}]$.

The axioms and rules of inference of Intuitionistic Quantifier Logic (often called IQC in the literature) are used to derive results in $PrAn$; these are listed below for the readers convenience.

$\wedge$ int $A \to (B \to (A \wedge B))$

$\wedge$ elim $(A \wedge B) \to A$

$\wedge$ elim $(A \wedge B) \to B$

$\vee$ int $A \to (A \vee B)$

$\vee$ int $B \to (A \vee B)$

$\vee$ elim $(A \vee B) \to ((A \to C) \to ((B \to C) \to C))$

$\to$ Ax1 $A \to (B \to A)$

$\to$ Ax2 $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$

T int T

F elim $\perp \rightarrow A$

$\forall$ elim $\forall sA \rightarrow A[^t_s]$ (Provided term t is substitutable for s in A)

$\exists$ int $A[^t_s] \rightarrow \exists sA$ (Provided term t is substitutable for s in A)

MP $\frac{A \quad A\rightarrow B}{B}$

$\forall$ int $\frac{A\rightarrow B}{A\rightarrow\forall x[B]}$ (Provided x does not occur free in A)

$\exists$ elim $\frac{A\rightarrow B}{\exists x[A]\rightarrow B}$ (Provided x does not occur free in B)

The following equality axioms are taken to hold in $PrAn$.

$PrAn$-$E1$ $\ulcorner t_1 = t_1 \urcorner$ where t is any term.

$PrAn$-$E2$ $\ulcorner t_1 = t_2 \urcorner \wedge \ulcorner t_1 = t_3 \urcorner \rightarrow \ulcorner t_2 = t_3 \urcorner$ where t_1, t_2 and t_3 are like terms.

$PrAn$-$E3$ $\ulcorner t_1 = t_2 \urcorner \iff \ulcorner S(t_1) = S(t_2) \urcorner$ where t_1 and t_2 are either both numerical or extended numerical terms.

$PrAn$-$E4$ $\ulcorner t_1 = t_2 \urcorner \rightarrow \ulcorner T(t_1, s_1, s_2, ..., s_i) = T(t_2, s_1, s_2, ..., s_i) \urcorner \wedge \ulcorner T(s_1, t_1, s_2, ..., s_i) = T(s_1, t_2, s_2, ..., s_i) \urcorner \wedge ... \wedge \ulcorner T(s_1, s_2, ..., s_i, t_1) = T(s_1, s_2, ..., s_i, t_2) \urcorner$ where T is a functional term and $s_1, ..., s_i, t_1$ and t_2 are either numerical or extended numerical terms.

$PrAn$-$E5$ $\forall x[S(x) \neq 0]$ where $x \neq y$ abbreviates $x = y \rightarrow \perp$.

The following axioms are taken to hold for $\prec$ in $PrAn$.

$PrAn$-$Or1$ $\ulcorner \nabla \prec t \urcorner$ where t is a numerical term.

$PrAn$-$Or2$ $\ulcorner t \prec \nabla \urcorner$ where t is a numerical term.

$PrAn$-$Or3$ $\ulcorner \nabla \prec \Delta \urcorner$

Both induction and primitive recursive closure are taken to hold in $PrAn$ and axioms for these are given below.

PrAn-Ind $A(0) \wedge \forall x[A(x) \rightarrow A(S(x)] \rightarrow \forall x[A(x)]$

PrAn-PRC $\exists\mu\forall x[\mu(x) = t(x)]$, where t is any term in the language that does not contain μ.

We introduce the following useful pieces of meta notation; we will write (assuming all function variables mentioned are of type $N \mapsto N$),

$n * m$ to denote the natural number encoding the finite sequence $\langle n(0), n(1), ..., n(x), m(0), ..., m(x)\rangle$ via repeated operations of the pairing function p.

$\underline{n}$ to denote the natural number encoding the finite tuple of sequences $\underline{n}_0, ..., \underline{n}_x$ (via repeated pairing operations).

$\langle\rangle$ to denote the extended natural number encoding for the empty sequence (0) to make formulae more readable.

$x * n$ and $n * x$ to denote the natural numbers encoding the sequences $\langle x\rangle * n$ and $n * \langle x\rangle$ respectively.

$\bar{\mu}(x)$ to denote the the natural number variable n encoding the finite sequence $\langle\mu(0), \mu(1), ..., \mu(x)\rangle$.

$\langle\mu, \mu'\rangle$ to denote the function ν such that $\nu(x) = \begin{cases} \mu(\frac{x}{2}) \ if \ \exists y[x = 2 * y] \\ \mu'(\frac{x-1}{2}) \ otherwise \end{cases}$.

$n < m$ to denote $\exists n'[n * n' = m]$.

$n \subset \mu$ to denote $n(0) = \mu(0) \wedge n(1) = \mu(1) \wedge ... \wedge n(x) = \mu(x)$.

$\mu(\nu) = x$ to denote $\exists n \subset \nu[\mu(n) = x]$. Sometimes we shall simply write $\mu(\nu)$ when referring to such an x.

$\mu = \nu$ to denote $\forall x[\mu(x) = \nu(x)]$.

$\mu \mid \nu = \nu'$ to denote $\forall x\exists y[\mu(x * \bar{\nu}(y)) = \nu'(x) \wedge \forall z < y[\mu(x * \bar{\nu}(z)) = \nabla]]$, sometimes we will simply write $\mu \mid \nu$ directly when referring to such a ν'.

$x \in N$ to mean $x \neq \nabla \wedge x \neq \Delta$.

$\mu \in K_0$ to denote the following.

$\forall \nu \exists x[\mu(\bar{\nu}(x) \in N] \wedge \forall n \forall m[n < m \rightarrow (\mu(n) \prec \mu(m) \vee \mu(n) = \mu(m))]$ (i.e. μ satisfies $K_0 1$ and $K_0 2$ as defined in §2.2)

Given a formula A we write $A(x)$ to denote $A[^x_t]$, where t is the only numerical term in A.

This concludes the construction of $PrAn$. We will use $PrAn$ to construct all later systems to ease comparison between them.

In summary, $PrAn =$ Intuitionistic Predicate Logic $+PrAn$-$T1$-$T6 + PrAn$-$E1$-$E6 + PrAn$-$Or1$-$3 + PrAn$-$Ind + PrAn$-PRC and our intended interpretation for the language of $PrAn$ is natural numbers and (not necessarily lawlike) functions as in Draglin (1979).

§3.4. The System FIM

The system FIM (*Foundation of Intuitionistic Mathematics*) was first proposed in Kleene and Vesley (1965) as a system for intuitionistic analysis and is widely accepted as the first **formal** foundational system for intuitionistic analysis. Veldman (2008) presents an alternative foundational system to FIM, however the treatment of choice sequences is the same; they are considered purely extensional objects.

In intuitionistic logic many classical results are no longer true; to derive equivalent statements we require our system for analysis (FIM) to have some additional axioms added to $PrAn$. Specifically, a **continuity axiom** (which implements the ideas of continuity derived in §2.4), an **axiom of choice** and an **axiom of bar induction** (both of which allow us to derive powerful results such as the Uniform Continuity Theorem) are all taken to hold in FIM.

The axiom of continuity in FIM is given below.

$FIM\text{-}BC\text{-}N$ $\forall\mu\exists x[A(\mu,x)] \to \exists\nu \in K_0\forall\mu[A(\mu,\nu(\mu))]$

The axiom of choice in FIM is given below.

$FIM\text{-}AC\text{-}NC$ $\forall x\exists\mu[A(x,\mu)] \to \exists\mu'\forall x[A(x,(\mu')_x)]$

The axiom of bar induction in FIM is given below.

$FIM\text{-}BI_D$

$$\begin{aligned}&[\forall\mu\exists x[R(\bar{\mu}(x))]\\ &\wedge \forall n[R(n) \vee \neg R(n)]\\ &\wedge \forall n[R(n) \to A(n)]\\ &\wedge \forall n[\forall x(A(n * x)) \to A(n)]]\\ &\quad \to A(\langle\rangle)\end{aligned}$$

This concludes the construction of FIM, one of the earliest and arguably most general systems for the foundations of analysis.

To summarise, $FIM = PrAn + FIM\text{-}AC\text{-}NC + FIM\text{-}BI_D + FIM\text{-}BC\text{-}N$ and the intended interpretation of the language of FIM is the same as in $PrAn$.

We pause here to repeat an observation given in both Kleene and Vesley (1955) and Draglin (1988).

> ‘We originally supposed that [$AC\text{-}NC$] would be required in formalising Brouwer’s proof of his uniform continuity theorem ... But Vesley [in a later section] has managed using only [$AC\text{-}NN$]’ (Kleene and Vesley, 1955 pp.72-73)

> ‘In the theory FIM one can derive all the basic facts of Brouwer’s intuitionistic analysis, including the fan theorem and the theorem on the uniform continuity of real functions. Actually, in this connection, it suffices to use the more modest theory $PrAn + FIM\text{-}AC\text{-}NN + FIM\text{-}BI_D + FIM\text{-}BC\text{-}N$’ (Draglin, 1988 pp130)

This weaker system, we feel, is well worth defining as later (in §6.3) we show that the system we construct (in §5.1) is actually an extension of this weaker system when subject to a series of specific restrictions.

We define the system $FIM\text{-}AN = PrAn + FIM\text{-}AC\text{-}NN + FIM\text{-}BI_D + FIM\text{-}BC\text{-}N$ with the intended interpretation of the language being the same as in FIM.

§3.5. The Systems IDB, LS and CS

The systems IDB, LS and CS are attributed to Kreisel and Troelstra. We accredit these systems to both Kreisel and Troelstra (even though both published separate papers on the matter) because they worked together to form the final picture we see today, with both contributing in an equal measure. This section will give a brief recap of their notion of lawlessness before constructing the systems in the order IDB, LS and finally CS.

The formal notion of a lawless sequence first arose in Kreisel (1968), being defined as a choice sequence where 'the simplest kind of restriction on restrictions is made, namely some finite initial segment of values is prescribed, and, beyond this, no restriction is to be made' (Kreisel 1968). In this formulation of lawlessness [taken from Troelstra (1977), though we shall refer to as KT-LL, '*Kreisel and Troelstra Lawlessness'*, due to its genesis] a *lawless sequence* is a sequence where no intensional information (information about how it is constructed) is known (or all such information has been abstracted away) except for identity and a finite initial segment. An extreme special case was later included – that of a *proto-lawless* sequence which only has identity and no other intensional information at all (no initial segment). It is important to note that the fact lawless sequences are considered disjointed from lawlike sequences gives us a clear picture that Kreisel and Troelstra believed that lawlessness was a property defined by the way a sequence is given, be it the method of generation or the abstraction of information. In particular, this notion of lawlessness indicates that two lawless sequences may only have identical elements iff they are given to us in the same way; this idea forms the crux of one of Fletcher's paradoxes (see §3.7.2).

§3.5.1. The System IDB

We begin with $PrAn$ and make some slight changes to the symbols used in the language to reinforce our new intended interpretation of function variables (IDB is entirely lawlike). This will also help avoid confusion when we construct LS and CS from IDB. The new symbols we shall use to denote function variables are f and g.

We include axioms used to construct the set K of inductively defined neighbourhood functions (introduced in §2.6) as they form a crucial part of the theory IDB (we will use the meta notation e here for a constructive function on extended natural numbers encoding finite sequences).

$K1$ $\forall x[\lambda n.x \in K]$

$K2$ $\forall e[(e(\langle\rangle) = \nabla \wedge \forall x[\lambda n.e(x * n) \in K]) \to e \in K]$

$K3$ $\forall e[(e = \lambda n.x \vee (e(\langle\rangle) = \nabla \wedge \forall x[\lambda n.e(x * n) \in Q])) \to e \in Q] \to \forall e[e \in K \to e \in Q]$

IDB obtains a special status in the work of Troelstra and Kreisel as it will later be shown that choice sequences are eliminable in their works. We remind the reader that a continuity axiom is omitted because we only require those when dealing with choice sequences and these are not present in the language of IDB.

There is no axiom of bar induction in IDB; instead, the following axiom is used to obtain bar induction (see theorem 2.6.2).

$K4$ $K_0 \subseteq K$

The axiom of choice for IDB is as follows.

IDB-AC-NF $\forall x \exists f [A(x, f)] \to \exists g \forall x [A(x, (g)_x)]$, where $\ulcorner t(x) \urcorner$ is any numerical term in the language not containing f and containing only constructive parameters.

Finally we rewrite the axiom of primitive recursive closure of $PrAn$ in the language of IDB below.

IDB-PRC $\forall f \forall x [f(x) = t(x)]$, where $\ulcorner t(x) \urcorner$ is any numerical term in the language not containing f.

This concludes the construction of the system IDB, though it is worth pausing here before summarising to make an important point. IDB was constructed for the purpose of eliminating choice sequence variables, and we will see that the elimination theorems in LS and CS allow us to map results to the seemingly weaker IDB. While the tools created in IDB are entirely lawlike, their purpose is to help us explore lawless sequences and choice sequences in general.

In summary, $IDB = PrAn + K1\text{-}4 + IDB\text{-}AC\text{-}NF + IDB\text{-}PRC$ and the language of IDB is the same as that of $PrAn$; however, the intended interpretation is different as IDB was created to be entirely lawlike.

§3.5.2. The System LS

The first formal system utilising only lawless sequences was first proposed by Kreisel (1968) under the name L. This system contained an incorrect axiom that was corrected in Troelstra (1977) where the system was also renamed LS (*Lawless Sequences*). LS is a system that looks only at the extreme cases of choice sequences – lawlike and lawless sequences only. The fact that LS does not cover choice sequences in general is reflected in the notion of finite information utilised as we shall see. LS does not provide a foundation for analysis as it stands; it is only by constructing a system based on the projections of such lawless sequences that we actually arrive at a theory for analysis. However, many results in this projective system are derived from the axioms of LS.

We begin the construction of LS by extending the language of IDB to include variables for *lawless sequences* (α, β).

The term formation for LS is as in IDB with the following extensions.

$PrAn$-$T4$ is modified slightly to give the following.

LS-$T4$ Any function defined by Church's λ that does not contain **lawless sequence** parameters is a **functional** term.

The following term formation rules for lawless sequences are also included.

LS-$T8$ Any lawless sequence variable α is a **lawless** term.

LS-$T9$ Given a lawless term T and a numerical term t we have that $\ulcorner T(t) \urcorner$ is a **numerical** term.

Formula formation in LS follows the same rules as IDB except that we also allow lawless terms.

The following meta notational concepts are introduced; we write,

$\underline{\alpha}$ to represent a finite tuple of choice sequences, $\alpha_1, \alpha_2, ..., \alpha_x$.

$\neq (\alpha, \beta_1, ..., \beta_x)$ to abbreviate $\alpha \neq \beta_1 \wedge ... \wedge \alpha \neq \beta_x$.

$\#(\alpha_1, ..., \alpha_x)$ to abbreviate $\neq (\alpha_1, \alpha_2, ..., \alpha_x) \wedge \neq (\alpha_2, \alpha_1, \alpha_3, ..., \alpha_x) \wedge ... \wedge \neq (\alpha_x, \alpha_1, ..., \alpha_{x-1})$

$\underline{n} \subset \underline{\alpha}$ to abbreviate $\underline{n}_0 \subset \underline{\alpha}_0 \wedge \underline{n}_1 \subset \underline{\alpha}_1 \wedge ... \wedge \underline{n}_x \subset \underline{\alpha}_x$.

The following axioms governing lawless sequences are included in LS.

$LS1$ $\forall n \exists \alpha [n \subset \alpha]$ (Density)

$LS2$ $\forall \alpha \forall \beta [\alpha = \beta \vee \alpha \neq \beta]$ (Decidability)

$LS3$ $(A(\alpha, \underline{\beta}) \wedge \neq (\alpha, \underline{\beta})) \rightarrow \exists x [\forall \alpha' [(\neq (\alpha', \underline{\beta}) \wedge \bar{\alpha}(x) \subset \alpha') \rightarrow A(\alpha', \underline{\beta})]]$, where A contains no other free choice sequences. (*Open Data)*

$LS4$ $\forall\underline{\alpha}\exists f[\#(\underline{\alpha}) \rightarrow A(\underline{\alpha}, f)] \rightarrow \exists e \in K\forall\underline{n}[e(\underline{n}) \in N \rightarrow \exists g\forall\underline{\alpha}[\underline{n} \subset \underline{\alpha} \wedge \#(\underline{\alpha}) \rightarrow A(\underline{\alpha}, g)]]$, where A contains no other free choice sequences. (BC-C)

Our reason for this is that it is implied by the final formulation of $LS4$ (in earlier formulations the question was left open if $e \in K$ or merely K_0), since it has been adopted as K then this is a strong indicator that $K = K_0$ has been accepted over the 'extension principle' (see Van Der Hoeven 1982 pp.4 for more detail).

This concludes the construction of LS. The reader is invited to notice that LS has a schema of bar induction by virtue of the axiom $K_0 \subseteq K$ in IDB.

Theorem 3.5.2.1 : LS **Elimination**

There exists a mapping τ from the formulae of LS without free lawless variables onto the formulae of IDB such that,

1 $\tau(A)$ is A if A is a formula of IDB

2 $LS \vdash (A \iff \tau(A))$

3 $LS \vdash A$ iff $IDB \vdash \tau(A)$

A proof (and a mapping τ) can be found in Troelstra (1977). This theorem illustrates an interesting transformation from extensionally defined results about lawless sequences to results about the intensional lawlike sequences. Given that LS is supposed to extend IDB, finding that results from LS can be expressed in IDB is quite surprising. It essentially means that lawless sequences as taken by Kreisel and Troelstra do not bring anything new to intuitionistic mathematics; they are a 'figure of speech'.

To summarise, $LS = IDB +$ variables for choice sequences $+LS$-$T4 + LS$-$T8 + LS$-$T9 + LS1 + LS2 + LS3 + LS4$. Formulae in LS without free lawless variables can be expressed in IDB which shows that the addition of lawless sequences as understood by Kreisel and

Troelstra adds nothing new to IDB. The intended interpretation of the lawless sequence variables are the lawless sequences of Kreisel and Troelstra (see §3.2.2).

§3.5.3. The System CS

A system similar to CS was first proposed in Kreisel (1963) (called C) but, as with Kreisel's L, it was later refined by Troelstra (1977) with the addition of the axiom of analytic data. This refinement came at a high cost; previously C was valid over all choice sequences whereas the refined CS is not. A universe of choice sequences satisfying it (the so called GC sequences) was soon hypothesised in Troelstra (1977, appendix C), refined in Dummett (1977 §7.4) and further exposited in Van Der Hoeven (1982, chapters 2 and 3)). Unlike LS, the system CS does provide a foundation for analysis and a very good one; it satisfies all the criteria mentioned in Dummett (1977) – closure under continuous operations, strong continuity and strong choice.

We begin the construction of CS by extending the language of IDB to include variables for GC *sequences* (μ, ν).

The term formation for IDB (hence $PrAn$) is extended to include the following term formation rules for GC sequences.

CS-$T8$ Any choice sequence variable μ is a GC **sequence** term.

CS-$T9$ Given a GC sequence term T and numerical term t we have that $\ulcorner T(t) \urcorner$ is a **numerical** term.

Formula formation in CS follows the same rules as IDB except we also allow choice sequence terms.

The following meta notational concepts are introduced; we write,

$\underline{\mu}$ to represent a finite tuple of choice sequences, $\mu_1, \mu_2, ..., \mu_x$.

$\neq (\mu, \nu_1, \nu_2, ..., \nu_x)$ to abbreviate $\mu \neq \nu_1 \wedge \mu \neq \nu_2 \wedge ...\mu \neq \nu_x$.

$\#(\mu_1, ..., \mu_x)$ to abbreviate $\mu_1 \neq \mu_2 \wedge ... \wedge \mu_1 \neq \mu_x \wedge \mu_2 \neq \mu_3 \wedge ... \wedge \mu_{x-1} \neq \mu_x . \#\mu_0, ...\mu_i$

$n \subset \underline{\alpha}$ to abbreviate $n \subset \underline{\alpha}_0 \wedge n \subset \underline{\alpha}_1 \wedge ... \wedge n \subset \underline{\alpha}_x$.

We then introduce the following axiom of primitive recursive closure.

CS-PRC $\exists\mu\forall x[\mu(x) = t(x)]$ where $\ulcorner t(x) \urcorner$ is any numerical term in the language not containing μ.

The Axioms for GC sequences are as follows.

CS-GC1a $\forall e \in K \forall\mu\exists\nu[e \mid \mu = \nu]$ (closure under continuous operations)

CS-GC1b $\forall\mu\forall\nu\exists\mu'[p(\mu, \nu) = \nu']$ (closure under pairing)

CS-GC2 $A(\mu) \to \exists e \in K[\exists\nu[\mu = e \mid \nu] \wedge \forall\nu[(e \mid \nu)]]$ (analytic data)

CS-GC3 $\forall\mu\exists f[A(\mu, f)] \to \exists e \in K_0 \forall n[e(n) \in N \to \exists f \forall\mu[n \subset \mu \to A(\mu, f)]]$ (continuity)

CS-GC4 $\forall\mu\exists\nu[A(\mu, \nu) \to \forall\mu\exists e \in K_0[A(\mu, e \mid \mu)]]$ (local function continuity)

The axiom of choice in IDB is strengthened in the following way.

CS-AC-NC $\forall x\exists\mu[A(x, \mu)] \to \exists\gamma\forall x[A(x, (\gamma)_x]$

This concludes the construction of CS. Before summarising, we will note (without proof) two important theorems in CS, BC-C and the CS elimination theorem and prove that bar induction holds in CS.

Theorem 3.5.3.1: Strong Continuity

$\forall\mu\exists\nu[A(\mu, \nu)] \to \exists e\forall\mu[A(\mu, e \mid \mu)]$ (BC-C)

For a proof please see Troelstra and Dalen (1988 vol2 pp.671).

Theorem 3.5.3.2 : CS Elimination

There exists a mapping τ from the formulae of CS without free choice sequence variables onto the formulae of IDB such that:

1 $\tau(A)$ is A if A is a formula of IDB.

2 $CS \vdash (A \iff \tau(A))$

3 $CS \vdash A$ iff $IDB \vdash \tau(A)$

A proof (and a mapping τ) can be found in Troelstra (1977). This theorem, while similarly stated to the one in LS, gives us a different insight. This mapping takes results about choice sequences (some without the lawlessness property) to results about lawlike sequences. So, while we may find it easier to work with the choice sequences, they do not give us any new results in this system, and again we must conclude that it seems they add nothing to the theory IDB.

In summary, $CS = IDB +$ variables for arbitrary choice sequences $+ CS\text{-}T8 + CS\text{-}T9 + CS\text{-}PRC + CS\text{-}AC\text{-}NC + CS\text{-}GC1\text{-}4$. Formulae in CS without free lawless variables can be expressed in IDB thus showing that choice sequences as envisioned by Kreisel and Troelstra bring nothing new to IDB. The intended interpretation of choice sequence variables are the GC sequences of Kreisel and Troelstra (see §3.2.5).

The next section will construct the systems of Moschovakis so that we can later compare her work with that of Kreisel and Troelstra.

§3.6. The Systems $FIRM\text{-}INT$ and $FIRM$

The systems of Moschovakis are largely of a classical flavour; the main goals of her work being to produce a common foundational system containing foundational subsystems for both intuitionistic and classical analysis. The formal systems of this section first appeared in Moschovakis (1987), later being developed and refined throughout Moschovakis (1993, 1994, 2016) .

Our main interest lies with her most recent intuitionistic foundational system that we shall call $FIRM\text{-}INT$ as Moschovakis never named this system [though it is very similar to $IDLS$

as expounded in Moschovakis (1993)]. Moschovakis, instead, defined $FIRM$-INT as a subsystem of a classically (though not intuitionistically) valid joint foundational super-system $FIRM$ [first expounded in Moschovakis (2016)]. We will construct these two systems in reverse order, first producing the intuitionistically valid subsystem $FIRM$-INT and then adding in the additional axioms to produce the classical joint foundational system $FIRM$. For the sake of notational consistency, we will provide minor notational modifications where appropriate to make comparison with the foundational systems introduced in the previous subsection easier.

§3.6.1. The System $FIRM$-INT

The system $FIRM$-INT forms the foundational system for Moschovakis; all axioms therein may be argued for intuitionistically if we take one slight leap of faith (which we shall address later). We construct it as follows.

In $PrAn$ we have general function (which may be seen as choice sequences as in FIM) variables already; to these we add an extra pair of function variables for 'lawlike' functions – f and g.

Term formation is extended as follows.

FI-$T2a$ Any function variable f is a **functional** term.

FI-$T6$ Any functional term t not containing free choice sequence variables is an **R-functional** term.

FI-$T7$ Any numerical variable t not containing free choice sequence variables is an **R-numerical** term.

Any formula containing only R-numerical and R-functional terms is an R-formula.

The following rules of predicate logic are introduced especially for R-functional terms.

*FI-LOG*1 $\frac{A \to B(f)}{A \to \forall f[B(f)]}$, if f is not free in A.

*FI-LOG*2 $\frac{\forall f[A(f)]}{A(t}$, if t is an R-functional term free in A.

*FI-LOG*3 $\frac{A(t)}{\exists f[A(f)]}$, if t is an R-functional term free in A.

*FI-LOG*4 $\frac{A(f) \to B}{\exists f[A(f)] \to B}$, if f is not free in B.

We next define a predicate on choice sequences RLS which forms the mainstay of Moschovakis' notion of lawlessness.

*FI-RLS-P*1 $\forall \mu[RLS(\mu) \iff \forall f[\forall n \exists m[f(n) = m] \to \exists x[\bar{\mu}(x) * f(\bar{\mu}(x)) \subset \mu]]]$

'μ is 'Relatively Lawless' iff every function on finite sequences will, at some point, 'predict' at least one term of μ'. We pause to note here that the statement $\forall n \exists m[f(n) = m]$ indicates that f maps natural number representing finite sequences to natural numbers representing finite sequences; this is not to be misconstrued as Moschovakis' notion of lawlike functions including partial functions.

This notion of relative lawlessness will be explored in more detail in the following section (§3.7); for now we ask the reader to accept it so that we may complete the definition of Moschovakis' systems before we subject them to any analysis. For notational convenience we define the class of relatively lawless choice sequences as follows.

$\mu \in M_{RLS} \iff RLS(\mu)$

Another core notion in the theory is that of two sequences may be lawless relative to one another, which is denoted by $RLS(\mu, \nu)$ and defined below.

*FI-RLS-P*2 $\forall \mu \forall \nu[RLS(\mu, \nu) \iff RLS(\langle \mu, \nu \rangle)]$

The reader is invited to note that $RLS(\mu) \wedge RLS(\nu) \to RLS(\mu, \nu)$, but the converse is not true!

We define a formula as *restricted* iff it has choice sequence quantifiers only in subformulae of the form listed below.

$$\forall\mu_0[RLS(\langle\mu_0,\mu_1,....,\mu_i\rangle)\rightarrow A] \textbf{ or } \forall\mu_0[RLS(\langle\mu_0,\mu_1,....,\mu_i\rangle)\wedge A]$$

We, next, present four axioms specifically for RLS choice sequences below.

$FI\text{-}RLS1$ $\forall n\exists\mu\in M_{RLS}[n\subset\mu]$

$FI\text{-}RLS2$ $\forall\mu\in M_{RLS}\forall n\exists\nu_{RLS(\langle\mu,\nu\rangle)}[n\subset\nu]$

$FI\text{-}RLS3$ $\forall\mu\in M_{RLS}[A(\mu)\rightarrow\exists x\forall\nu\in M_{RLS}[\bar{\mu}(x)\subset\nu\rightarrow A(\nu)]]$, where A is restricted and has no other free choice sequence variables.

$FI\text{-}RLS4$ $\forall\mu\in M_{RLS}\exists f[A(\mu,f)]\rightarrow\exists g\exists f\forall\mu\in M_{RLS}[\exists!x[g(\bar{\mu}(x))\in N]\wedge\forall x[g(\bar{\mu}(x))\in N\rightarrow A(\mu,\lambda y.f(\langle g(\bar{\mu}(x)),y\rangle))]]$, where A is restricted with no other free choice sequence variables.

The next step is to strengthen the axiom of choice as described below.

$FI\text{-}AC\text{-}NC$ $\forall x\exists\mu[A(x,\mu)]\rightarrow\exists\nu\forall x[A(x,(\nu)_x]$

Next we add in an axiom of bar induction (the axiom originally listed in her system) as follows.

BI_T

$$\begin{aligned}&\forall\mu\exists!x[R(\bar{\mu}(x))]\\&\wedge\forall n[R(n)\rightarrow A(n)]\\&\wedge\forall n[\forall z[A(n*z)]\rightarrow A(n)]\\&\rightarrow A(\langle\rangle)\end{aligned}$$

The following schema of primitive recursive closure is also included.

$FI\text{-}PRC$ $\forall f\forall x[f(x)=\ulcorner t(x)\urcorner]$ where $\ulcorner t(x)\urcorner$ is any numerical term in the language not containing f.

We introduce below the final axiom in $FIRM\text{-}INT$, the axiom of general continuous choice.

$FI\text{-}GC$ $\forall\mu[A(\mu) \to \exists\nu[B(\mu,\nu)]] \to \exists\gamma\forall\mu[A(\mu) \to E(\gamma \mid \mu) \wedge B[\mu, \gamma \mid \mu]]$, where A contains no $\vee$ or $\exists$ except immediately before a formula between terms, and $E(\nu)$ is read as 'ν is well defined' (see Troelstra and Dalen 1988 pp.13 for more details on this predicate).

This concludes the construction of $FIRM\text{-}INT$; to summarise, $FIRM\text{-}INT$ can be stated as $PrAn+FI\text{-}T2a+FI\text{-}T6+FI\text{-}T7+FI\text{-}AC\text{-}NC+BI_T+FI\text{-}RLS1-4+FI\text{-}GC+FI\text{-}PRC$. The language of $FIRM\text{-}INT$ is the language of $PrAn$ with additional variables for 'lawlike' sequences. The intended interpretation of the lawlike sequence variables is the same as in IDB.

§3.6.2. The System $FIRM$

To construct the system $FIRM$ (valid neither intuitionistically or classically) we, first, extend the language of $FIRM\text{-}INT$ to include the predicate symbol $\ll$ ['the lawlike sequences are well ordered by $\ll$' (Moschovakis (unpublished)].

Formula formation is extended to allow $\ulcorner t \ll s \urcorner$, where t and s are functional terms.

Six axioms defining the binary relation $\ll$ are given below.

$FIRM\text{-}W0$ $(\mu = \nu \wedge \mu \ll \nu' \to \nu \ll \nu') \wedge (\nu = \nu' \wedge \mu \ll \nu \to \mu \ll \nu')$

$FIRM\text{-}W1$ $\forall f \forall g(f \ll g \to \neg g \ll f)$

$FIRM\text{-}W2$ $\forall f \forall g \forall g'(f \ll g \wedge g \ll g' \to f \ll g')$

$FIRM\text{-}W3$ $\forall f \forall g(f \ll g \vee f = g \vee g \ll f)$

$FIRM\text{-}W4$ $\forall f(\forall g(g \ll f \to A(g)) \to A(f)) \to \forall f A(f)$ where A is any R-formula

$FIRM\text{-}W5$ $\mu \ll \nu \to \neg\forall f \forall g \neg(\mu = f \wedge \nu = g)$

Moschovakis (2016) states that 'the double negation in W5 is essential', her reason being its use in her realisability model for the system. Also worthy of note is that $W5$ would restrict the ordering to the lawlike sequences without the enforcement of this double negation.

To this we add the restricted law of excluded middle given below.

$FIRM\text{-}RLEM$ $\forall\mu \in M_{RLS}[A(\mu) \vee \neg A(\mu)]$, where A is restricted and with no other arbitrary choice sequence variables free.

This concludes the construction of $FIRM$; to summarise, $FIRM = FIRM\text{-}INT + FIRM\text{-}RLEM + FIRM\text{-}W0-5$ with the language of $FIRM\text{-}INT$ extended to include the predicate symbol $\ll$. The intended interpretation of the language of $FIRM$ is as in $FIRM\text{-}INT$ if one wishes to consider it intuitionistically.

§3.6.3. Closed Data

A reader familiar with the work of Moschovakis may question why we do not include her axiom of 'closed data' in any of the systems presented. This brief subsection aims to illuminate the reasoning for this.

The axiom of closed data [actually a theorem under the classical interpretation according to Moschovakis (2016)] is given in the language of $FIRM\text{-}INT$ as follows.

CD $\forall\mu \in M_{RLS}[\forall y \exists \nu \in M_{RLS}[\bar{\mu}(y) \subset \nu \wedge A(\nu)] \rightarrow A(\mu)]$

'If we can find, for every initial segment of μ, a choice sequence ν sharing that initial segment which satisfies $A(\nu)$ then we can conclude $A(\mu)$'.

Moschovakis (2016) states that 'In an intuitionistic subsystem obtained by omitting [$FIRM\text{-}RELM$], [CD] may be taken as an additional axiom schema' which rather leaves its inclusion as a decision to the reader. However in the similar system $IDLS$, Moschovakis (1994) states that 'The last axiom of [the subsystem of $IDLS$], ... , is the principle of closed data', leaving one to question why its requirement for inclusion was dropped, though communication via e-mail with Moschovakis cleared this up quickly and confirmed that the schema is optional and may be left out in our intuitionistic subsystem.

The intuitionistic validity of this axiom is in serious doubt as, given the definition of lawlessness expounded by Moschovakis, when intuitionistically read the $\forall y \exists \nu$ quantifier combination raises a problem as seen below.

> If for any y we could construct a ν such that $\bar{\mu}(y) \subset \nu$ then this implies we have a lawlike way to predict the elements of μ via our very ability to construct such ν's from y, and hence, surely, μ could not be lawless in the sense of Moschovakis; in essence, CD is vacuous when read intuitionistically!

This is not to say that the schema is not valid and of use classically; of course, merely that it adds nothing to the theory for an intuitionist. It is for this reason that we decline to include this optional schema in the system we shall explore.

§**3.7. Comparing** LS **With** $FIRM\text{-}INT$

In this subsection we will first introduce Fletcher's paradoxes in §3.7.1; after this we will then compare the notions of Kreisel and Troelstra lawlessness ($KT\text{-}LL$) and *Moschovakis' lawlessness* ($M\text{-}LL$) in §3.7.2 before finally comparing LS with $FIRM\text{-}INT$ directly in §3.7.3.

§**3.7.1. Fletcher's Paradoxes**

Fletcher (1998 pp.132–135) brings to light several paradoxes (collectively known as *Fletcher's paradoxes)* which arise with some of the current theories of choice sequences, also giving some indication as to why such paradoxes arise. These will prove important for comparing $FIRM\text{-}INT$ with LS, as well as helping shape our own theory later in this text. Fletcher's motivation for constructing these paradoxes was to demonstrate his belief that 'classifying the choice sequences into proto-[KT-LL], [KT-LL] and non-[KT-LL] is a mistake' (Fletcher, 1998). These paradoxes were originally formulated to examine the notion of KT-LL, and thus we shall preserve this intention as we state them below.

$FP1$ Given that μ is lawless, barring the constraint that each term must be 0 or 1, define the sequence ν as follows: $\forall x[\nu(x) = 1 - \mu(x)]$. From this set up, μ is lawless but ν is not. Both cannot be lawless as this violates open data; but surely ν also has a reasonable claim to lawlessness?

The counter that Troelstra is likely to make is based on a similar construct set out in Troelstra (1977 pp.16 and pp.48) and is that 'we cannot refer to [μ] and [ν] as both being lawless within the same context'. Fletcher quickly points out that Dummett disagrees; 'it is not a matter of there being two choice sequences, either of which we may, if we will, take as being a lawless sequence, provided that we do not so take the other, but of this being impossible we should know [the mirror image relation] to hold of any two lawless sequences' (Dummett, 1977 pp.421). From this disagreement on such a fundamental point Fletcher concludes 'The mirror-image example throws serious doubt on the notion of a [KT-LL] sequence' (Fletcher, 1998 pp.133) and we are strongly inclined to agree with this conclusion.

$FP2$ Given any lawless sequence μ and define ν to be the sequence obtained by prefixing μ with 0 as the first term. We have that μ is lawless and thus, so must ν be. However, this violates open data once again.

$FP3$ Given the proto-lawless sequence (see §3.2.2 for a definition of proto-lawless sequences) μ define ν to be the lawless sequence obtained by replacing the first term of μ by x. From this set up, if we assume $\mu = \nu \iff \mu \equiv \nu$ [which Troelstra clearly does for lawless sequences (see Troelstra 1977 pp.14)], we can obtain the result that no proto-lawless sequence has a first element (Fletcher, 1998 pp.133-134).

$FP4$ Given the proto-lawless sequence μ, define the (assumed lawless) sequence ν as μ with the first 10 terms removed. Both of these sequences are related by a continuous operation, but $\mu \not\equiv \nu$. ν cannot be proto-lawless as that would violate open data; yet it meets all the criteria to be proto-lawless!

Fletcher's conclusion from these paradoxes is that the conventional explanations of KT-LL 'do not adequately justify the principles that are attributed to them and the theory that is based on them' (Fletcher, 1998 pp.135). The notion of KT-LL is problematic in and of itself, though now the question arises as to how it compares with M-LL (the notion of lawlessness defined in terms of the predicates $RLS1$ and $RLS2$ in §3.6.1).

§3.7.2. Comparing Notions of Lawlessness

The notions of KT-LL and M-LL are quite different, both in their philosophical outlook and their mathematical structure.

KT-LL is considered to be a constructed intensional property of a sequence; a property that is inbuilt in some way or deliberately added via some conceptual abstraction operator. To be concise, KT-LL is a second order restriction (a restriction on restrictions) on the choice sequence in question. This notion of lawlessness is easier for an intuitionist to accept as all it involves is the imposition of some form of second order restriction, although Brouwer' rejection of second order restrictions (see §3.2.1 for more detail) does stir unease.

M-LL is a very different take on lawlessness; it is an extensional notion which is defined via a predicate ($RLS1$ or $RLS2$). It is entirely extensional in that it relies on checking every possible 'predictor' to ensure that it successfully 'predicts' at least one element of the choice sequence under scrutiny; there is no second order restriction involved. A clear refinement over KT-LL is the notion that lawless sequences may be related in some way and still remain lawless ($RLS(\mu) \wedge RLS(\nu)$ does not preclude there being some function mapping μ to ν or vice versa), though we still have a tool that allows us to enforce independence where it is necessary (the two placed RLS predicate). Regrettably, even in light of these refinements, we find ourselves a little uncomfortable as intuitionists since both versions of the RLS predicate are clearly undecidable so the intuitionistic justification for the existence of such a M-LL sequence is still very much an open question.

In essence, M-LL requires a leap of faith to be viable (the existence of at least one M-LL choice sequence) whereas KT-LL is immediately much more palatable to an intuitionist, though an advocate of Brouwerian methods may well remain uneasy in the face of the second order restriction required for lawlessness.

When exposed to Fletcher's paradoxes the notion of M-LL fairs better than KT-LL as seen below.

$FP1$ and $FP2$ are no longer valid as a result of the refinements of M-LL over KT-LL. Moschovakis allows the interlacing of sequences and, thus, predicates over two sequences are constructed in the guise of a predicate over one sequence with the strong restriction that the interlacing itself must satisfy RLS. In both of the cases explored by Fletcher the interlaced sequence formed does indeed satisfy this property.

$FP3$ and $FP4$ are no longer valid as Moschovakis has no notion of proto-lawlessness sequence and both these paradoxes rely on the difference between lawless and proto-lawless sequences.

Overall, M-LL fairs far better in light of Fletcher's paradoxes than KT-LL; however, M-LL requires a leap of faith that many intuitionists would be loath to make. In essence, neither notion of lawlessness is perfect conceptually and thus some further refinement is desirable.

§3.7.3. Comparing the Systems LS and $FIRM$-INT

Moschovakis (1987) compares one of her earlier systems with LS. While her final system differs slightly from its earlier incarnation, much of this comparison is still valid and any comments made by Moschovakis in this section will be cited. To make reading easier, we will write α and β for KT-LL choice sequences and μ, ν for general (not necessarily lawless in either sense) choice sequences.

In $FIRM$-INT and LS we have 4 axioms for choice sequences, though the content of these axioms is quite different. We will now proceed to compare the single choice sequence cases of these axioms side by side.

$LS1$ $\forall n \exists \alpha (n \subset \alpha)$

FI-$RLS1$ $\forall n \exists \mu \in M_{RLS}[n \subset \mu]$

$LS1$ and FI-$RLS1$ actually share a similar meaning, though the notion that α has some intensionally lawless property and μ is a general choice sequence satisfying some extensional property of lawlessness makes a significant difference. FI-$RLS1$ is harder to swallow for an intuitionist since it relies on an undecidable predicate.

$LS2$ $\forall \alpha \forall \beta (\alpha = \beta \vee \alpha \neq \beta)$

$LS2$ has no analogue in $FIRM$-INT, having its roots in the strong assumption of Kreisel and Troelstra that extensional equality and intensional equality coincide. In $FIRM$-INT, $LS2$ is omitted simply because there is 'no reason to assert it' (Moschovakis 1987). It is from this axiom that $FP3$ springs, which is why FI-$RLS3$ is not valid for M-LL.

FI-$RLS2$ $\forall \mu \in M_{RLS} \forall n \exists \nu [RLS(\langle \mu, \nu \rangle) \rightarrow n \subset \nu]$

FI-$RLS2$ is introduced since M-LL allows two M-LL sequences to be related and this axiom ensures us that entirely independent M-LL choice sequences do in fact exist; it is, in essence, a second density axiom allowing us to construct these independent choice sequences for any given initial segment.

$LS3$ $(A(\alpha, \underline{\beta}) \wedge \neq (\alpha, \underline{\beta})) \rightarrow \exists x[\forall \alpha'[(\neq (\alpha', \underline{\beta}) \wedge \bar{\alpha}(x) \subset \alpha') \rightarrow A(\alpha', \underline{\beta})]]$, where A contains no other free choice sequences.

$FI\text{-}RLS3$ $\forall \mu \in M_{RLS}[A(\mu) \rightarrow \exists x \forall \nu \in M_{RLS}[\bar{\mu}(x) \subset \nu \rightarrow A(\nu)]]$, where A is restricted and has no other free choice sequence variables.

$LS3$ and $FI\text{-}RLS3$ both express open data and, while they look very similar, there is actually a stark difference; the sequences in $LS3$ must all be be apart; whereas, the single sequence given in $FI\text{-}RLS3$ may be comprised of other interlaced $M\text{-}LL$ sequences and no such apartness is enforced.

$LS4$ $\forall \alpha \exists f[A(\alpha, f)] \rightarrow \exists e \in K \forall n[e(n) \in N \rightarrow \exists f \forall \alpha_{n \subset \alpha}[A(\alpha, f)]]$, where A has no other free choice sequence variables

$FI\text{-}RLS4$ $\forall \mu \in M_{RLS} \exists f[A(\mu, f)] \rightarrow \exists g \exists f \forall \mu \in M_{RLS}[\exists! x[g(\bar{\mu}(x)) \in N] \wedge \forall x[g(\bar{\mu}(x)) \in N \rightarrow A(\mu, \lambda y.f(\langle g(\bar{\mu}(x)), y \rangle))]]$, where A is restricted with no other free choice sequence variables

These axioms show stark differences as seen below.

Firstly, the $\exists e \in K$ is very different to the $\exists g[\forall \mu \in RLS \exists! y[e(\bar{\mu}(y)) \in N]]$. The latter, we feel, being the weaker of the two.

Secondly, the power of the additional $\exists f$; in $LS4$ it is $\forall n \exists f$ whereas in $FI\text{-}RLS4$ it is $\exists f \forall n$, a notably stronger statement!

This makes their direct comparison difficult, though according to Moschovakis 'The analogue for $[FIRM\text{-}INT]$ of the continuity principle $LS4$, for restricted A, is likewise derivable from $[FI\text{-}RLS4]$; but the analogue for LS of $[FI\text{-}RLS4]$ is simpler and apparently weaker than $LS4$' (Moschovakis 1987).

The other differences between $FIRM\text{-}INT$ and LS stem from the fact that $FIRM\text{-}INT$ has variables for **general** (and lawlike) choice sequences, whereas, LS only has variables for **lawless** (and lawlike) choice sequences. This, of course, is to be expected due to the different notions of lawlessness both systems embrace.

This concludes the comparison of LS and $FIRM\text{-}INT$. To summarise, we have found that both systems have their flaws; LS is prey to paradox when considered informally, does not seem to add any new results (see the comments on the elimination theorems) and is not a viable foundation for analysis when taken alone; $FIRM\text{-}INT$ requires a great deal of faith on the part of intuitionist reader due to the definition of $M\text{-}LL$ and its density notion relies on an argument that intuitionists would find hard to swallow.

§3.8. Ideas of the Creative Subject

The theory of the creative subject (also known as 'the creating subject') is not a formal system for choice sequences; it is a collection of rules about knowledge and understanding which form the basis for Brouwer's strong counter-examples. One may wonder why we include it at all if it does not pertain directly to a formal theory on choice sequences; the author argues that given the matter brought to light by Niekus' work, a reappraisal of the ideas offered by the theory of the creative subject may provide fresh material for those seeking an alternative notion of a choice sequence suitable for founding analysis.

We will begin by informally introducing the notion of the creative subject and stating its purpose. Following this, we shall provide the standard interpretations and axioms for the creative subject before presenting a strong non-standard alternative given by Niekus.

§3.8.1. Origins of the Creative Subject and its Nature

The theory of the creative subject finds its genesis in Brouwer (1948) (which Brouwer actually called the 'creating subject'). Brouwer's notion of the creative subject was deliberately vague; he never clearly states the rules or axioms that the creative subject were bound by. The path to the formalisation of the creative subject can be found in Kreisel (1967) with refinements in Myhill (1968) and Troelstra (1969) for those who are curious, though the reader is warned that it was developed as 'separate' from the systems explored above.

In each set of axioms, the notion $\vdash_x A$ was used to mean 'at time x the *creative subject* (usually read as 'some idealised mathematician') has a proof of A'. What it means for us to 'have a proof of A' is still widely debated and, as such, we shall devote §3.8.2 to exploring the three interpretations presented by Dummett of what is meant. §3.8.3 will be devoted to the axioms of the creative subject as well as a notion known as *Kripke's Schema*. Finally, in §3.8.4, we will explore a much more recent discourse on the creative subject by Niekus arguing that Brouwer really meant a choice sequence.

§3.8.2. Standard Interpretations

Dummett (1977) lays down three alternative interpretations of what $\vdash_x A$ means – the *strict*, *intermediate* and *lenient* interpretations. While Dummett states that:

> 'Probably the most sensible attitude is that it is indeterminate which of these interpretations of '$\vdash_x A$' should be adopted, and that we are free to choose between them' (Dummett, 1977 pp.343).

he was quick to note that:

> 'There are, however, difficulties both about the lenient and intermediate interpretations' (Dummett, 1977 pp.343).

We present our own arguments for the superiority of the strict interpretation below.

Strict ‘*It* may be demanded that we have explicitly recognised the statement as having been proved, that is, we have not only effected a construction constituting a proof of it, but are consciously aware that we have done so’ (Dummett, 1977 pp.342).

This interpretation requires not only that the creative subject has the proof at hand but also that it is recognised as a proof. This interpretation certainly ‘feels’ the most constructive and thus appears the most attractive. It is commonly maligned due to a paradox attributed to Troelstra (1969 pp.105–107). However Dummett (1977 pp.346) appears to offer a satisfactory resolution to this. Accepting this interpretation is not without its advantages as it gives access to a stronger set of creative subject axioms (see Dummett 1977 pp.340–342 for more detail).

Intermediate ‘*We* have proved a statement just in case we have effected a construction which would, by itself, be a proof of that statement, whether or not we have noticed that it is so’ (Dummett, 1977 pp.343).

This interpretation removes the requirement that we understand that the construct we have is a proof of the point. The trouble is that if we do not recognise the proof then how would we be in any position to assert $\vdash_x A$ at all? Surely, if we did not recognise what we have as a proof of A, we would make no claims to it being a proof of A? It seems almost contrary to basic intuitionistic thought to assert otherwise. There is an interesting mirror case here, namely the case where we recognise what would be considered a proof, and we know how to construct such a proof, but we have not yet done so. Unfortunately, this case falls into a similar trap; how can we recognise something as a proof and assert that we can provide such a proof without having first having provided such a proof?

Lenient ‘We may be regarded as having proved a statement provided that we have explicitly proved one or more statements from which it follows very directly’ (Dummett, 1977 pp.343).

To assert $\vdash_x A$ all one requires is $\vdash_x A'$ (interpreted strictly or intermediately) for some A' (or collection of) from which a proof of A 'follows very directly', a term that is in no way clearly defined. To quote Martino (1982) there is a 'difficulty in generally characterising in a sufficiently precise manner the vague notion of [very directly]' or using the original point made by Dummett (1977 pp. 334) 'the lenient interpretation is afflicted with the problem that attaches to all attempts to distinguish between immediate and remote consequence'. This interpretation does not require a proof and it is very ambiguous as to whether we must recognise that our conclusion follows from our explicitly proved 'antecedents'. The requirement that this 'following very directly' be recognised is not present, though it seems reasonable to assume it is to gain any coherent notion from this interpretation. This interpretation, in essence, feels as if we are simply allowing ourselves to miss out 'obvious' steps in proofs where the definition of 'obvious' is in question. Again, accepting this interpretation is not without its advantages as it allows access to even stronger axiom schema than the intermediate interpretation.

§3.8.3. Standard Creative Subject Axioms

The following axioms are the ones presented in Dummett (1977) and seem to be representative of those present in all the literature. We give the axioms along with a reading based on the interpretations in the previous subsection to illustrate just how much changes based on one's interpretation. The reader is advised that the 'we' in the readings refers to the creative subject and not ourselves. This is a convention introduced by Troelstra (1969) that has endured in the literature to this day.

$CSu1$ $\forall x(\vdash_x A \vee \neg \vdash_x A)$

'At any time index x we can **recognise** whether or not we have proven A' (strict).

'At any time index x we either **"have"** a proof of A or we do not **"have"** a proof of A' (intermediate).

'At any time index x we either **recognise/"have"** the "direct antecedents" for a proof of A or we do not **recognise/"have"** the direct "antecedents" for a proof of A' (lenient).

$CSu2$ $\forall x \forall y[\vdash_x A \rightarrow \vdash_{x+y} A]$

'Once we have proven something it remains proved' (all).

$CSu3$ $\exists x \vdash_x A \rightarrow A$

'If we **recognise** a proof of A at time x then we can transform this into a proof of A' (strict).

'If we **"have"** a proof of A at time x then we can transform this into a proof of A' (intermediate).

'If we **recognise/"have"** the "direct antecedents" for a proof of A at time x then we transform this into a proof of assert A' (lenient).

$CSu4$ $A \rightarrow \neg\neg\exists x[\vdash_x A]$

'If we have a proof of A then we can transform this into a proof stating its not the case that there we will not be able to "find" some time where we **recognise** a proof of A' (strict).

'If we have a proof of A then we can transform this into a proof stating its not the case that there we will not be able to "find" a time where we **"have"** a proof of A' (intermediate).

'If we have a proof of A then we can transform this into a proof stating its not the case that there we will not be able to "find" a time where we **recognise/"have"** the proofs for a series of "direct antecedents" for a proof of A' (lenient).

It is sometimes preferred to assert a stronger notion of 4 which we shall call $4*$.

$CSu4*$ $A \to \exists x[\vdash_x A]$

> 'If we have a proof of A then we can transform this into a proof stating we must be able to "find" a time where we **recognise** a proof of A' (strict).
>
> 'If have a proof of A then we can transform this into a proof stating we must be able to "find" a time where we **"have"** a proof of A' (intermediate)
>
> 'If we have a proof of A then we can transform this into a proof stating we must be able to "find" a time where we **recognise/"have"** the proofs for a series of "direct antecedents" for a proof of A'

As can be seen above, the interpretation of $\vdash_x A$ is exceedingly important, with the meaning of certain axioms changing entirely. On the whole, it is the author's opinion that the strict interpretation is the closest to the notion of intuitionistic thought, the main reasons for this being given below.

1. The notion of possessing a proof and not recognising it is not dubious in itself; however, the claim that we can state that we 'have' a proof without recognising it indicates that someone or something must have recognised this proof to make such a claim. In essence we find ourselves at the mercy of a 'superior' idealised mathematician who can recognise results that our idealised mathematician can not, a notion that goes against the whole idea of idealisation. The mirror image of this case is also unsatisfactory for similar reasons.

2. The notion of recognising the 'antecedents' for a proof without knowing the proof itself is inherently suspicious; how would you recognise them without also recognising how they led to a proof of the statements? This matter becomes worse if we interpret things even more liberally and do not assume we recognise these proofs.

3. Who is actually doing the proving? So far we have been following the assumption that the 'we' in all readings is actually some idealised mathematician, i.e. some entity other than ourselves. Brouwer himself never clarifies this matter and most of our readings are based on the interpretation given in Troelstra (1969). It is possible that this interpretation might not be the correct one (§3.8.4 explores an alternative reading where 'I' is used).

Regardless of how we read these axioms, we are able to obtain a useful tool known as Kripke's Schema (first published in Myhill 1966), which allows us to construct Brouwer's weak counter-examples from them via the function $\mu(x) = \begin{cases} 0 \; if \; \neg(\vdash_x A) \\ 1 \; if \; \vdash_x A \end{cases}$. We give both the weak (assumed axioms 1,2,3 and 4) and strong (assumed axioms 1,2,3 and 4*) versions of *Kripke's Schema* below.

KS $\exists\mu[\forall x \forall y[x \leq y \to \mu(x) \leq \mu(y) \leq 1] \wedge \forall x[\mu(x) = 1 \to A]$

KS+ $\exists\mu[\forall x \forall y[x \leq y \to \mu(x) \leq \mu(y) \leq 1] \wedge (\exists x[\mu(x) = 1] \iff A)]$

§3.8.4. Niekus' Interpretation

In Niekus (1987) we find an entirely different interpretation as to what $\vdash_x A$ means, as well as an original notion proposed in section 3 of Niekus (1987) that suggests that Brouwer was simply using choice sequences, and that any reference to a 'creating subject' is just a figure of speech. These ideas were developed further in Niekus (2005 and 2010), both of which we shall draw upon to present a clear notion of the ideas expressed by Niekus' work.

Niekus originally argued for a different interpretation of $\vdash_x A$ than the strict interpretation (he does not even consider the lenient or intermediate interpretations), citing Troelstra's paradox as his reason for rejecting the strict interpretation. This interpretation is best stated in Niekus (2010) and we give this version below.

> 'We imagine our future to be covered by a discrete sequence of ω stages, starting with the present stage as stage 0, and we define for a mathematical assertion

$[\vdash_x A]$

as: at the x^{th} stage from now we shall have a proof of $[A]$.' Niekus (2010).

As can be seen, this notion makes use of tensed logic 'will have' rather than 'has'. More importantly, Niekus (2017) specifically insists that 'Since in Brouwer's view mathematics is a creation of the human individual, the expression "creating subject" may mean nothing but just "the mathematician". So it can be "I" or "we". We shall use "we" for [the creative subject]'. In essence, Niekus is denying the need for a notion of an idealised mathematician to make use of his interpretation.

We present below the conventional creating subject axioms read in this way.

$CSu1$ $\forall x(\vdash_x A \vee \neg \vdash_x A)$

'At any time index x we either know we **will have** proven A or **will have** not proven A'.

$CSu2$ $\forall x \forall y[\vdash_x A \rightarrow \vdash_{x+y} A]$

'Once something has been proven it will remain proven'.

$CSu3$ $\exists x \vdash_x A \rightarrow A$

'If we know we **will have** a proof of A at time x then this can be transferred to a proof of A'.

$CSu4$ $A \rightarrow \neg\neg\exists x[\vdash_x A]$

'If we have a proof of A then this is sufficient to say that it is not the case that we will not be able to 'find' some time where we know a proof of A'.

It is sometimes preferred to assert a stronger notion of 4 which we shall call $4*$ and present below.

$CSu4*$ $A \rightarrow \exists x[\vdash_x A]$

> 'If we have a proof of A then this allows us to prove that we must be able to 'find' a time where we will know a proof of A'.

Something important that requires mention is that Niekus does not specify if we would recognise our proofs, but merely that we can construct them. Under the argument laid down in §3.8.2 it seems more likely that we would have to recognise such proofs since you or I would not be able to assert that something will be a proof of something else without recognising it as a proof. This is a key difference since we can not make this inference under the conventional interpretation (since the idealised mathematician need not be human).

Niekus proves (in Niekus 2010) that, under his interpretation, $CSu1$ and $CSu3$ are false, and hence that Kripke's Schema is also not generally valid under this interpretation, though he is quick to add that 'there are one or more instances of [Kripke's Schema] for specific cases in the work of Brouwer, he always avoided its use as a general principle for an unspecified formula' (Niekus, 2010). Though almost to counter this point a footnote on the same page of the same work says 'There is an instance of [Kripke's Schema] in Brouwer's work, from the last year in which he published, see [Collected works pp.252]. Whether there are arguments for this specific instance of [Kripke's Schema] remains an interesting question'. Given that even the weakest form of continuity (WC-N) is not consistent with Kripke's Schema (see §2.4.2), we find this as a point in favour of his argument for this interpretation over those outlined in Dummett (1977).

As a result of this interpretation, Niekus (1987, 2005 and 2010) argues that Brouwer's writings were misinterpreted and, rather than introducing some new notion, Brouwer was simply making use of choice sequences in his arguments. As Niekus (2010) concludes at the end of the fourth section,

> 'We claim that Brouwer is using an incomplete sequence ... It is not the introduction of an idealised mathematician that makes the creating subject arguments special, but the

application of individual incomplete objects.' (Niekus, 2010)

where 'individual incomplete objects' refers to choice sequences. This seems to tacitly imply that choice sequences have properties analogous to axioms 2, 4 and 4* – a notion we find no contradiction with. Overall, Niekus offers a very real and convincing interpretation to the mystery of the creative subject, one which, if accepted, resolves a central conflict in the theory of the foundations of intuitionistic analysis.

§3.9. Literature Review Conclusion

Before we offer a summary, we first feel compelled to mention a recent system devised in Kachapova (2015) [which in turn is a refinement of the system put forth in Bernini (1976)] that contains axioms for the creating subject (along with lawless sequences and axioms similar to those laid down by LS, though with $LS4$ missing). This system introduces 'higher types' of function; however, it is the author's opinion that this adds nothing new to the theory; such objects are already included via the simple trick of interlacing sequences or pairing natural numbers in FIM. This system admits both weak continuity and Kripke's schema which places its consistency in a very questionable state given the incompatibility of these two schemata.

This chapter has shown that there are a number of conflicts within the existing literature – be it the definition of lawlessness, the notion of what a choice sequence is or even the use of time within the formal theory.

The central critiques of most interest to us of the existing literature are given below.

1. The concept of lawless sequences as a whole is flawed in its current state; neither of the main interpretations given by Moschovakis and Troelstra offer a path free of philosophical difficulty.

2. Systems such as CS and FIM provide a satisfactory foundation for analysis; however, even these cannot agree on what role choice sequences play in the foundations of analysis, be it as a 'figure of speech' or central objects. More importantly, the complexity inherent in the notion of GC sequences for (the only universe of choice sequences constructed so far that satisfy CS) and the pure extensionality of FIM leave us wondering if there is not an alternative route one may take that sheds more light on the foundations of intuitionistic analysis.

3. The theory of the creative subject is a notion too rooted in time and, as Niekus points out, may be the result of a misinterpretation of Brouwer's ideas.

4. No theory fully explores a stronger notion of finite information than that offered by initial segments which, given the precarious position of weak continuity, seems an unusual oversight.

The concept of finite information explored in the next chapter aims to provide a framework to remedy at least 1 and 4, the resolution of 2 requiring further work.

We end this section with figure 2 over-leaf; outlining all of the systems we have constructed in this chapter.

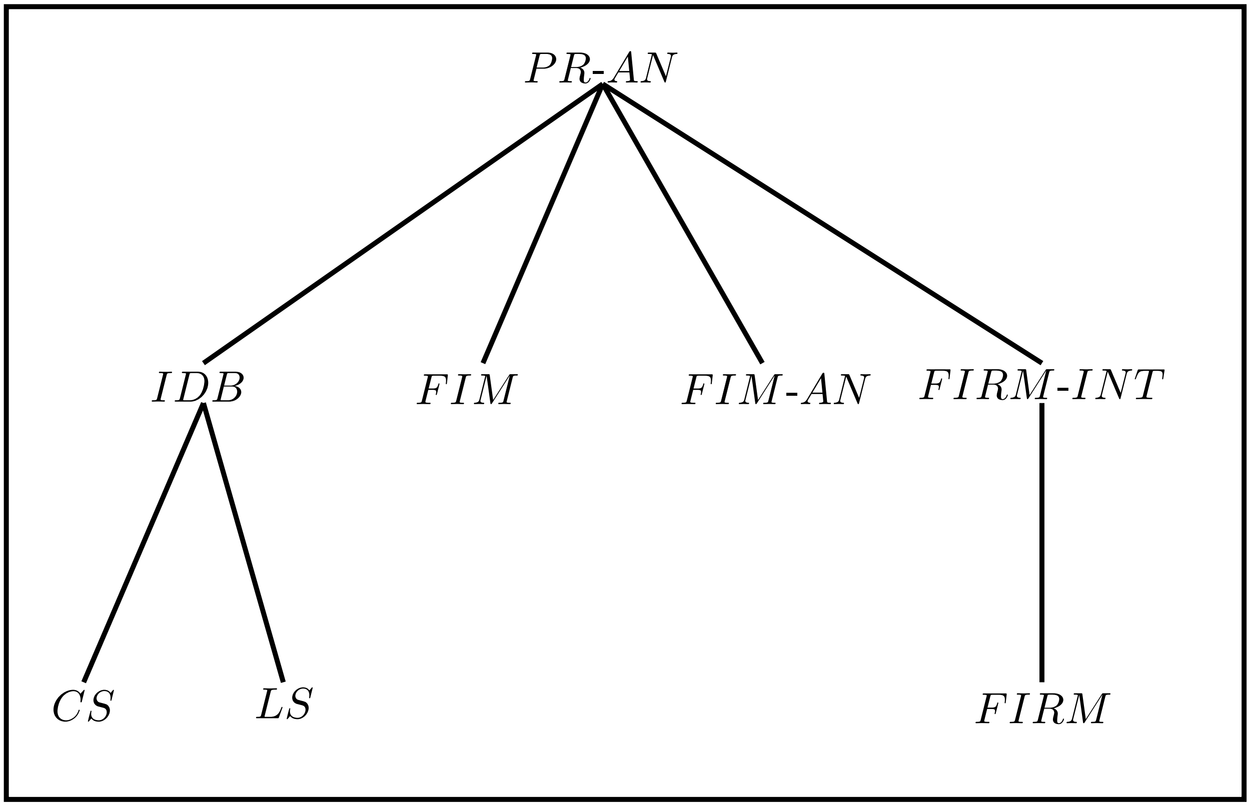


Fig 2 : Existing systems

4. The Theory of Knowledge States

§4.1. Outline

In this chapter we put forward our generalisation of the notion of 'finite information', provide generalised versions of conventional axioms and attempt to provide generalised proofs for common theorems in the existing theory. In the previous chapters we have seen the same method being followed with regards to finite information; all information save initial segments is abstracted away. The theory of knowledge states is the author's own attempt to refine this and provide a language with which we can express additional forms of intensional information about choice sequences; as such, it forms the bulk of original material in this thesis. Our reasons for wanting to do this are simple; the way we handle additional finite information lies at the heart of Fletcher's Paradoxes and, thus, at the heart of the ills plaguing the intuitionistic notion of lawlessness. We also find ourselves forced to make a great many tacit assumptions in our path to analysis and it is the author's hope that a path with fewer assumptions may one day be forged. The theory of knowledge states succeeds in producing a more coherent (though not perfect) notion of lawlessness, as well as allowing us to present an alternative theory containing less assumptions; one which may pave the way to analysis without so many restrictions.

The following outline lists the main aim of each section: §4.2 will informally outline our notion of knowledge states and their relations, it will also introduce our two 'axioms of knowledge' and prove generalised versions of open data and weak continuity. §4.3 will provides a formalisation for knowledge states about a single choice sequences; §4.4 will then extend this formalisation to knowledge states about tuples of sequences. §4.5 introduces an ordering on knowledge states, defines a useful form of equality on knowledge states, and shows that the ordering of knowledge states is a lattice. §4.6 explores some central notions for analysis, namely: contradictory knowledge states; the 'one manual rule'; our notion of 'lawlessness' and, it defines two key species of knowledge state. §4.7 introduces our notion of 'knowledge continuity' and proceeds to define a species of functions ($\hat{K}_0$) that act as our neighbourhood functions. §4.8 defines a generalised version of BC-N and explores the derivation of AC-NN from this. §4.9 explores how our system fares with Fletcher's Paradoxes. Finally, §4.10 explores a modified notion of bar induction in the language of knowledge states.

§4.2. Choice Sequences and Knowledge States

We shall call a never-ending process continuously generating natural numbers the *generating process* of a choice sequence. Some examples of generating processes are: tossing a coin and writing down the outcome; a constructive function from N to N, and a man writing down the number of stars he sees in the sky each night.

We shall call any intensional information about a generating process (even one that proves impossible to carry out), or finite list of numbers produced by such a generating process, a *knowledge state.* We shall reserve the symbol σ to denote knowledge states as needed in formal notation.

We say that *a knowledge state σ is consistent with the choice sequence μ* when σ is one of the following,

1. Part of the definition of the generating process of μ.
2. A list of elements observed from the generating process of μ.
3. A combination of both of these.

We denote 'σ is consistent with μ' by the notation $\sigma(\mu)$.

Our language of knowledge states will treat the extensionally derived elements of a choice sequence and the intensionally defined information about a choice sequence in the same way. We will explore the idea of the intensional information within knowledge states with the three examples below.

Example 4.2.1

Let us say we have two choice sequences μ and ν such that some knowledge state consistent with μ tells us that it is generated by the law $\lambda x.2x$ and some knowledge state consistent with ν tells us that it is generated by the law $\lambda x.\frac{4x}{2}$. These two sequences satisfy $\forall x[\mu(x) = \nu(x)]$, so we can assert that these sequences are **extensionally the same** ($\mu = \nu$), but we cannot be sure if they are **intensionally distinct** ($\mu \not\equiv \nu$) or if they are **intensionally the same** ($\mu \equiv \nu$). This is because our knowledge states are only 'part of the definition of the generating process[es]' of μ and ν. Whether this 'part' forms the whole or not is entirely unknown.

Example 4.2.2

Let us now say we have the choice sequence μ such that some knowledge state consistent with μ tells us that it is generated by the law $\lambda x.2x$ and that it is generated by the law $\lambda x.\frac{4x}{2}$. This knowledge state offers two different methods of generation for the same sequence which may seem paradoxical, since surely a sequence can only have one method of generation? I would argue against this since here we are not only asserting that μ is generated by both of these methods but that we are also affirming that these different methods of generation are consistent (will always provide the same values) as such information can only be read directly from the generating process of a choice sequence (this tacit assumption that the object we are given is a choice sequence is the only assumption we need make, which seems reasonable enough). If this is the case then it surely **does not matter** which one is used.

Example 4.2.3

Finally, let us consider two knowledge states; one indicating that a sequence belongs to the universal spread, the other indicating that it belongs to the spread s defined as follows $s(n) = 0 \iff \forall x < lth(n)[n(x) \; is \; not \; an \; odd \; perfect \; number]$. Are these two knowledge states extensionally equivalent? To say that they are (or are not) would be to say that we have resolved the existence (or non existence) of an odd perfect number, which is of course not the case! Thus, we are forced to conclude that extensional equality between knowledge states will not always be decidable.

The key points from these three examples are summarised below.

1. Knowledge states are treated intensionally.
2. Extensional equivalence between knowledge states is undecidable.
3. Knowledge states do not always fully define the generating process of a choice sequence.
4. Given any law generating the elements of a choice sequence, there may be other extensionally equivalent laws that are not defined in the generation process which also generate those elements.
5. The generating process of a choice sequence may contain more than one law.
6. Extensional equivalence of two knowledge states is not decidable in general
7. We can not infer intensional identity between two choice sequences via one knowledge state about each alone.

Conclusions 6 and 7 do not prove problematic as we never rely on extensional equivalence of knowledge states, nor on discovering intensional equivalence of choice sequences in our formal theory for analysis. These conclusions, instead, lead us to believe that this theory is closer to the notion of natural thought and discovery; in life, we may have two objects which appear the same that may, in essence, be very different, and we may never know if this is the case.

It is simple to construct a knowledge state that does not describe the generating process of any choice sequence, i.e. one that describes a process that is impossible to execute. A simple example is the knowledge state that indicates to us that all of the elements of a choice sequence are in the spread of even numbers **and** that all of the elements of the same choice sequence are in the spread of odd numbers. We do not attempt to ban such *inconsistent knowledge states* as this would be impossible due to the undecidability of extensional equivalence mentioned above. Instead, we make use of the fact that when we have a knowledge state consistent with a choice sequence, we know this because we have either obtained it directly from observing the form of the generating process or from observing the generating process in action (or both).

We introduce an interesting example below as it highlights a borderline case that needs clarification before we may continue.

Example 4.2.4

> Given two choice sequences μ and ν, a knowledge state consistent with μ that tells us it is generated by the law $\lambda x.2x$, and a knowledge state consistent with ν that tells us it is generated by both $\lambda x.2x$ and $\lambda x.\frac{4x}{2}$.
>
> Does our knowledge about ν allow us to make any inferences about the laws governing μ? If we just consider μ alone, then, looking at example 4.2.1, we seemingly cannot make the inference that μ is generated by $\lambda x.\frac{4x}{2}$ and $\lambda x.2x$.
>
> The intensional knowledge state consistent with ν demands that these laws be consistent (always give the same output) via the very fact that ν is taken to be a choice sequence.
>
> This proof should be sufficient to 'bridge the gap' and hence, yes, we can assert that $\lambda x.\frac{4x}{2}$ generates the elements of μ; however, we cannot assert that it is part of the intensional generating process of μ. So we may say there is some relation between the elements of μ and ν (extensional equality), but we may not assert that the $\lambda x.\frac{4x}{2}$ is explicitly mentioned in the generating process of μ.

This example shows us just how strongly intensional our treatment of knowledge states is; the only information we can use is the information in the knowledge state itself. We cannot use the information that both $\lambda x.2x$ and $\lambda x.\frac{4x}{2}$ generate ν (and thus must be extensionally equal) to imply that $\lambda x.\frac{4x}{2}$ also generates ν. In general, we cannot use intensional information about one sequence to produce new intensional information about another choice sequence. This idea will prove to be vital when we formulate our refined notion of lawlessness later on.

Example 4.2.4 also allows us to confirm an important fact – the decidability of $\sigma(\mu)$ for any σ and any μ. Given any σ and any μ, because of how we define consistency and because of our strong intensionality requirements, we can always verify if $\sigma(\mu)$ or $\neg\sigma(\mu)$, i.e. consistency between a choice sequence and a knowledge state is always decidable **as long as** we assume that all knowledge about μ is available to us. This is not something we shall generally assume until we begin our construction of analysis (§4.2.1 will go into this in more detail); hence, we do **not** assert that consistency between a choice sequence and a knowledge state is decidable in general **but** it will be when we begin our reduction to analysis.

Continuing onward, we remind the reader that M denotes the species of all choice sequences. As we develop the formal theory it will also be handy to discuss the species of all knowledge states, the species of all choice sequences consistent with some σ and the species of all knowledge states consistent with some μ. For notational simplicity, we shall denote these by Σ, M_σ and Σ_μ respectively.

We pause to give an additional definition (which we make no use of, however it is of sufficient philosophical import to be worth mentioning); one could define extensional equality between knowledge states as the following.

$$\sigma = \sigma' \iff M_\sigma = M_{\sigma'}$$

As our system makes no use of extensional equality of knowledge states this notion will not be used, and is only stated for the reader's interest.

We now extend our notation to cover the case of a tuple of sequences $\mu_0, \mu_1, \mu_2, ..., \mu_{x-1}$, which we will denote as $\underline{\mu}$ for notational convenience. To refer to a specific choice sequence in a tuple $\underline{\mu}$ we use its index, so $\underline{\mu}_4$ refers to the 5^{th} sequence of the tuple. To denote the number of choice sequences contained in a tuple we write $|\underline{\mu}|$, so $|\mu_0, \mu_1, ..., \mu_{x-1}| = x$.

We denote the number of sequences σ refers to by $|\sigma|$, so $|\sigma| = 2$ would mean σ is a knowledge state containing specific information about two choice sequences.

We write σ *is consistent with* $\mu_1, \mu_2, \mu_3, ..., \mu_x$ as $\sigma(\mu_1, \mu_2, \mu_3, ..., \mu_x)$ and take this to mean that one of the following is true.

1. σ defines part of some of the generating processes of the choice sequences $\mu_1, \mu_2, \mu_3, ..., \mu_x$.
2. σ contains elements of the sequences $\mu_1, \mu_2, \mu_3, ..., \mu_x$ obtained by observing their generating processes.
3. σ is some combination of both of the above.

It would be grammatically inconsistent to have a knowledge state about two choice sequences consistent with a single choice sequence ($\sigma(\mu)$ where $|\sigma| = 2$). The converse case ($|\underline{\mu}| > |\sigma|$) seems reasonable, since if we have a knowledge state which informs us that 'the first element of the first sequence is 1' then by saying it is consistent with two sequences ($\sigma(\mu, \nu)$) already implies that 'we know nothing about the other sequence', a statement that has no impact on the validity of our consistency. Attempting to prevent this case would also prevent a clear definition of order in knowledge states (defined in §4.6). To outline these two crucial ideas we define the following axiom.

AX-MOD $\forall\sigma\forall\underline{\mu}[\sigma(\underline{\mu}) \rightarrow |\sigma| \leq |\underline{\mu}|\ ]$.

From now on, when we have any tuples of objects mentioned (for example $\forall\underline{\mu}\exists\underline{\nu}$), we will assume that they are tuples of the same size unless otherwise mentioned. The restrictions on any knowledge state σ mentioned in such quantifiers will be as outlined in the axiom above.

When speaking about a tuple of unnamed objects we naturally tend to index them, for example, if talking about a tuple of unnamed cats we would use the phrases 'the first cat, the second cat, etc' (in essence, we give them a name) and I see no reason our knowledge states should not reflect this nuance of language. Given a knowledge state about two choice sequences we will refer to these as the first sequence and the second sequence. For example, $\sigma(\mu, \nu)$ would place μ as the first sequence and ν as the second sequence. This idea extends to any finite tuple of choice sequences; so $\sigma(\underline{\mu})$ would refer to $\underline{\mu}_1$ as the first sequence, $\underline{\mu}_2$ as the second sequence, etc.

We offer the following example to help clarify this notation.

Example 4.2.5

> Let us now say we are given a μ and a ν such that some knowledge state consistent with both sequences states that the first sequence is generated by the law $\lambda x.2x$, and the second sequence is related to the first in that it is generated by adding one to every element of the first sequence. If we had $\sigma(\mu, \nu)$ then μ would be generated by the law $\lambda x.2x$ and ν would be generated by the relation $\lambda x.\mu(x) + 1$. On the other hand, if we had $\sigma(\nu, \mu)$ then ν would be generated by the law $\lambda x.2x$ and μ would be generated by the relation $\lambda x.\nu(x) + 1$.

The reader is invited to notice that $\sigma(\mu, \nu)$ **contradicts** $\sigma(\nu, \mu)$!

At this point, it will be useful to look more closely at the relationships between choice sequences, knowledge states consistent with choice sequences and the elements of choice sequences. A helpful idea is to picture the choice sequence (its generating process) as a black box labelled with its identity. The elements of this sequence are printed out one by one (in no specific order) and are stored in a pile next to it. So far, this idea coincides with the notion of 'black boxes' introduced in Fletcher (1998 pp.135) to eliminate the temporal component of choice sequences; however, we extend this model slightly by modelling our intensional information about this sequence as a manual containing information about this black box. Generally speaking, this manual need not contain all possible intensional information about the black box and our pile of elements most certainly will not contain every possible element of a choice sequence. So we can query this black box and obtain any element we like (and add to our pile), and we can look in the manual to see what we know about the inner workings of this black box (fully knowing that the manual may be updated with more information later on).

A curious distinction that arises is that elements can be read from the black box and elements can be 'set by hand' in the manual. In the conventional theory, this distinction was maintained and become the difference between restrictions (elements 'set by hand' before generation began) and 'choosing' (elements 'generated' later), this is what separated the lawless and proto-lawless sequences. This distinction also introduced an unfortunate temporal component to the theory as 'restrictions' came **before** 'choices', some idea of time, however informal, was required to fully understand the distinction. In our model, the idea of 'elements set by hand' are elements printed in the manual and 'generated elements' are elements in our printout pile.

To prevent this temporal problem from occurring in our theory, we abstract away the 'source' of any element, ignoring how we came about it. We can do this because if we obtain an element from querying our black box then this is sufficient 'proof' to 'pencil it into the manual' so to speak. It is guaranteed that the information we 'pencil in' is consistent with all other information contained within the manual (since it was obtained from the box itself and the box makes Fletcher's invariance assumption); in essence, we treat all knowledge as if it were intensional.

The example below illustrates this idea in practice.

Example 4.2.6

Given some μ generated by the law $\lambda x.x + 1$, our current knowledge about μ is just its generating law. Let us say we query μ for its first element and get $\mu(0) = 1$. We can now assert some more knowledge about μ, namely that its first element is 1, because for $\mu(0)$ to be 1 this must be a result of the generating process of μ. We essentially convert our discovered element into some intensional fact about the generating process of μ and assume we knew it all along. So now we have two knowledge states about the same sequence, one stating the generating law and another stating that the first element is 1. Their consistency is guaranteed and can easily be checked since, as a unique property of dealing with elements 'set by hand', we can always check consistency.

In our theory, something of great importance is that we can always query the black box for **any** element at **any** time and it will always give it to us. Knowledge about specific elements is special in this way; we can always be certain of obtaining information about any desired element of a choice sequence.

§4.2.1. A Brief Digression in the Interests of Analysis

So far in our examples, when we have spoken of black boxes and their manual, we have been tacitly assuming that each black box has **only one** manual that is extended in some way when we learn something new. A notion closer to Brouwer's concept of finite information (outlined in §3.2.1) would be that we have a collection of manuals that are more informative the 'later' they are written, a notion that preserves some form of temporal element which we wish to avoid when working formally with choice sequences, although it is perfectly useful at a conceptual level. These differences aside, there is still some temporal notion remaining, and we wish to propose a schema here that both eliminates this time and still grants us a universe of choice sequences suitable for analysis. The schema proposed in this section is **not** something we enforce before we come to our reduction before analysis; however, it is best introduced here so that all further material can be read in light of it by the more utilitarian mathematician if so desired. In essence, it introduces the idea of some maximal (intensional) knowledge state for a choice sequence, a notion that grants us some useful properties listed later in this section.

The idea of a maximal knowledge state directly contradicts Brouwer's notion of a choice sequence (see §3.2.1); such a maximal state could not exist in his interpretation since we could have a sequence with an additional restriction added at each stage forever. Restrictions on restrictions must be static in Brouwer's theory (though perhaps not in Dalen and Atten's); otherwise we could never be certain whether a choice sequence were lawless or not since it might turn out to be hesitant (starts out KT-LL and then adopts a law at some juncture) instead! However, all we are really doing is placing an additional second order restriction in place – that our 'manual' is always the definitive one. This notion is, surely, no stronger than that demanded by lawless choice sequences of Kreisel and Troelstra (see §3.2.2 for what is essentially their 'no manual' rule); indeed, it is a far gentler restriction than the one imposed by lawless choice sequences.

It has already been noted that we treat information about specific elements differently than all other kinds of information. The way in which we shall treat such information shall remain quasi-temporal; this temporal component is easily abstracted away as the argument in the previous subsection showed. For the purpose of producing a notion of finite information suitable for analysis we informally define a **'one manual' rule**, i.e. given any choice sequence μ, there exists some 'maximal' knowledge state consistent with it that contains **all** the intensional information about that choice sequence. Full formalisation will be given in §4.6, however, we present the core idea we intend to formalise below.

> Given any choice sequence, there exists some unique knowledge state containing **all** intensional information (all information that isn't finite lists of specific sequence elements)

about that choice sequence; any other knowledge state containing intensional information about that choice sequence will be 'contained' within this knowledge state.

An important point to note is that two tuples of sequences $\underline{\mu}$ and $\underline{\nu}$ might share the same maximal intensional knowledge state yet still be extensionally distinct. A trivial example being two sequences μ and ν such that the maximal knowledge state for both is $Spread(s_{uni})$ where s_{uni} denotes the universal spread. That is, when talking of a maximal knowledge state, we are excluding additional intensional information that isn't in Σ_{SE}, we treat finite lists of elements differently than other intensional information. Thus two sequences sharing the same maximal knowledge state is not always sufficient to even ensure their extensional equivalence, let alone their intensional equivalence, as their extensional information may differ.

Once we have a formal language of knowledge states, and a formal concept of order on knowledge states defined, we will be able to offer a fully formalised definition for this notion

To continue using our metaphor of a manual, we are saying that, at some point, the manual may only be further updated with SE atoms. This rule seems to contradict what we previously stated in the previous section (that the manual could always possible be extended) and, indeed, this would be the case if we were intending to use the entire universe of choice sequences to obtain our notion of analysis (which we are not). The one manual rule essentially allows us to define a sub-universe of choice sequences for the purpose of analysis. That this sub-universe is of sufficient cardinality to give us analysis is beyond question, since the universe of lawless sequences is contained within the universe of sequences with a single manual (the blank manual); indeed, our universe is far larger as the notion of lawlessness we will later define is far more generous than the existing one.

This '*one manual rule*' is essentially a meta-restriction preventing the growth of our finite information in a specific way and this does have some unusual effects as it:

1. Allows us to assert intensional identity between *Knowledge State Lawlike* (KS-LE)

choice sequences (sequences whose unary knowledge state contains a law).

2. Allows us to construct *Knowledge State Lawless* (KS-LL) choice sequences (choice sequences where naught but elements are contained in its unary knowledge states) without the need for any additional second order restrictions.

3. Renders consistency between knowledge states and choice sequences as decidable.

4. Ensures that given any choice sequence, we can always construct an extensionally equal KS-LL choice sequence.

5. Removes any potential residual temporal element that remained in our system.

The reason we take this path for analysis is simply because it is the path of least resistance; it prevents us having to define some complex abstraction operation to construct finite information for lawless choice sequences and, also, neatly 'snips off' any residual temporal notions we may be tempted to infer.

For now, though, we shall **not** assume the 'one manual rule' and work instead with the full universe of choice sequences till we restrict our universe of choice sequences for the purpose of constructing analysis in §6.1.

§4.2.2. Resuming Choice Sequences and Knowledge States

With the relation between choice sequences and knowledge states clarified, we may begin to consider how this relates to predicates on choice sequences and tuples of choice sequences.

Given any predicate A and a tuple of valid variables for that predicate, it is intuitionistically valid to say say that we either have a proof of A for that tuple of variables, have a proof of $\neg A$ for that tuple of variables, or we lack a proof asserting either instance for that tuple of variables. We do not know if we will ever find a proof of either A or $\neg A$ for that tuple of variables, though we do not rule it out.

If we have a proof of $A(\mu)$, then this proof must rely on some finite information about μ, since μ can never be completed, i.e. it relies on some knowledge state σ consistent with μ. We should thus be able to recognise the finite information required by this proof; otherwise, we would never be able to construct it in the first place!

For any predicate $A(\mu)$ on choice sequences, we devise a special **decidable** predicate $A'(\sigma)$ on knowledge states that checks if a knowledge state, σ, is sufficient to construct the proof of $A(\mu)$ for any μ such that $\sigma(\mu)$. The reason we demand on decidability of A' is that is its entire purpose; A' is a predicate we wish to be able to pass a knowledge state to and immediately determine if it contains sufficient information to prove A. Were it not decidable, there would be no purpose in introducing it as it would be entirely useless to us.

An important question to answer is 'given any A and μ can we always construct a viable A'?'. We shall refer to this as the *constructibility of A'*.

Given any predicate A, if we were to assert $A(\mu)$, then we would clearly recognise what is required for a proof of $A(\mu)$ since $A(\mu)$ is equivalent to saying we have, and also recognise, such a proof. In other words, A' would be constructible in this context. The trouble arises when we have a predicate A such that we are not sure if there is a μ satisfying it; here we are under no onus to insist that we have or recognise a proof for anything. Yet to insist that A' be constructible in this context would be to do just that. In essence, if we assume that A' is constructible for any A, we are forcing ourselves into a position that demands we know what is required to prove A before we state A. Worse yet, if A' were constructible for any A, we would be asserting that we have some method for proving A true for a given choice sequence, an idea that seems repellent. A potential solution would be to define A' as follows.

$$A'(\sigma) \iff \forall\mu[\sigma(\mu) \to A(\mu)]$$

However, this approach would mean A' is undecidable, a state of affairs that contradicts our initial demand that it be decidable. We shall, for now, err on the side of caution and **not** assume that A' is always constructible for any given A.

The following example highlights something we have so far avoided mentioning – additional choice sequences that may be required to prove $A(\mu)$.

Example 4.2.2.1

> Define A as follows: $A(\mu) \iff \exists\nu[\forall x[\mu(x) = \nu(x) + 1]]$. To prove that $A(\mu)$ holds knowing that $\mu(x) = \lambda x.\nu(x)+1$ would be sufficient, however (as we shall later see when we formalise such relations) this knowledge state is about two sequences, not one. In essence, the knowledge involved in the proof of a predicate **may** contain information about additional sequences.

In CS, LS, $FIRM\text{-}INT$ and FIM, we see no mention of such 'additional' sequences, though in these systems the reason behind this is simple. The authors of all these systems demand that the predicated contains 'no additional choice sequence variables'. We are not forced to take this path as our notion of finite information is far more robust.

With these ideas in mind, we can now begin to formulate a pair of axioms relating predicates over choice sequences and knowledge states that summarise the following pair of key ideas.

> Given that we have $A(\underline{\mu})$ for some arbitrary $\underline{\mu}$, then we only have it by virtue of some finite information about $\underline{\mu}$ and, possibly, some additional $\underline{\nu}$. More formally,

σ-μ-1 $\forall\underline{\mu}[A(\underline{\mu}) \rightarrow \exists\sigma[\exists\underline{\nu}[\sigma(\underline{\mu},\underline{\nu})] \wedge A'(\sigma)]]$, where A has no other free choice sequence variables and $|\underline{\mu}|$ need not equal $|\underline{\nu}|$.

> Given any σ such that $A'(\sigma)$ holds, then for any free choice sequences $\underline{\mu}$ (and possible some $\underline{\nu}$), we have that $\sigma(\underline{\mu})$ (or $\sigma(\underline{\mu},\underline{\nu})$) is sufficient for $A(\underline{\mu})$ to hold. More formally,

σ-μ-2 $\forall\sigma[A'(\sigma) \rightarrow \forall\underline{\mu}[\exists\underline{\nu}[\sigma(\underline{\mu},\underline{\nu})] \rightarrow A(\underline{\mu})]]$, where A has no other free choice sequence variables and $|\underline{\mu}|$ need not equal $|\underline{\nu}|$.

We pause here to comment on a potential 'higher order' knowledge state that would allow us to simplify our axioms. In this work we are only considering knowledge about a choice sequence's generating process; however it may well be possible to define a notion of knowledge about the knowledge consistent with a choice sequence. Something of the form $\forall \underline{\mu}[\sigma(\underline{\mu}) \iff \exists \underline{\nu}[\sigma'(\underline{\mu}, \underline{\nu})]]$ provides a perfect example. Such a notion would allow us to simplify our axioms of knowledge somewhat, however the addition of these additional forms of knowledge (in particular knowledge about knowledge) is fraught with potential complications (for if we allow knowledge about knowledge then we'd need to also define knowledge about knowledge about knowledge etc) and is better left as a subject for future work (see §7.3.2).

With these two axioms we can now prove two theorems – the converse of σ-μ-1 and a generalised form of Open Data. The converse of σ-μ-2 ($\forall \sigma[\forall \underline{\mu}[\exists \underline{\nu}[\sigma(\underline{\mu}, \underline{\nu})] \to A(\underline{\mu})] \to A'(\sigma)]$) is questionable at best, since it implies that every A has a unique A' which is demanding a lot from our A'! The author's attempts to construct a proof of it starting from σ-μ-1 have so far been unsuccessful.

Theorem 4.2.3.1 : Converse of σ-μ-1

$\forall \underline{\mu}[\exists \sigma[\exists \underline{\nu}[\sigma(\underline{\mu}, \underline{\nu})] \wedge A'(\sigma)] \to A(\underline{\mu})]$, where A has no other choice sequence variables and $|\underline{\mu}|$ need not equal $|\underline{\nu}|$.

Proof:

We begin by stating σ-μ-2.

$$\begin{aligned} & \forall \sigma[A'(\sigma) \to \forall \underline{\mu}[\exists \underline{\nu}[\sigma(\underline{\mu}, \underline{\nu})] \to A(\underline{\mu})]] \\ \iff & \forall \sigma \forall \underline{\mu}[A'(\sigma) \to (\exists \underline{\nu}[\sigma(\underline{\mu}, \underline{\nu})] \to A(\underline{\mu}))] \\ \iff & \forall \underline{\mu} \forall \sigma[(\exists \underline{\nu}[\sigma(\underline{\mu}, \underline{\nu})] \wedge A'(\sigma)) \to A(\underline{\mu})] \\ \iff & \forall \underline{\mu}[\exists \sigma[\exists \underline{\nu}[\sigma(\underline{\mu}, \underline{\nu})] \wedge A'(\sigma)] \to A(\underline{\mu})] \end{aligned}$$

as required. ♠

Theorem 4.2.3.2 : Generalised Open Data (GOD)

$\forall\underline{\mu}[A(\underline{\mu}) \rightarrow \exists\sigma[\exists\underline{\nu}[\sigma(\underline{\mu},\underline{\nu})] \wedge \forall\underline{\mu}'[\exists\underline{\nu}'[\sigma(\underline{\mu}',\underline{\nu}')] \rightarrow A(\underline{\mu}')]]]$, where A has no other choice sequence variables and $|\underline{\mu}|$ need not equal $|\underline{\nu}|$. (GOD)

Proof:

Given any $\underline{\mu}$ assume that $A(\underline{\mu})$ holds.

By σ-μ-1 we have that $\exists\sigma[\exists\underline{\nu}[\sigma(\underline{\mu},\underline{\nu})] \wedge A'(\sigma)]$.

By σ-μ-2 we can replace$A'(\sigma)$ with $\forall\underline{\mu}'[\exists\underline{\nu}'[\sigma(\underline{\mu}',\underline{\nu}')] \rightarrow A(\underline{\mu}')]$ to get $\exists\sigma[\exists\underline{\nu}[\sigma(\underline{\mu},\underline{\nu})] \wedge \forall\underline{\mu}'[\exists\underline{\nu}'[\sigma(\underline{\mu}',\underline{\nu}')] \rightarrow A(\underline{\mu}')]]$.

Discharging our hypothesis '$A(\underline{\mu})$' and quantifier '$\forall\underline{\mu}$' gives us the following.

$\forall\underline{\mu}(A(\underline{\mu}) \rightarrow \exists\sigma[\exists\underline{\nu}[\sigma(\underline{\mu},\underline{\nu})] \wedge \forall\underline{\mu}'[\exists\underline{\nu}'[\sigma(\underline{\mu}',\underline{\nu}')] \rightarrow A(\underline{\mu}')]])$ as required. ♠

This theorem is the most general form of open data possible. 'Given any property about choice sequences, it holds for any given choice sequence by virtue of the finite information we have about that choice sequence (and possibly some others)'. The fact that open data is a theorem in this theory and not an axiom is, I believe, a positive indicator that we are on the right track. It is now built into the way we understand finite information rather than imposed as an axiom.

Open data is all fine and well, but one of our key critiques of the conventional theory was the lack of any formalised justification of WC-N (see §2.4), and we have yet to do much better. The following theorem takes us beyond the conventional theory and provides us with a fully justified generalisation of WC-N.

Theorem 4.2.3.3 WC-N-KS

$\forall \underline{\mu}[\exists x[A(\underline{\mu}, x)] \rightarrow \exists \sigma \exists x' \forall \underline{\mu}'[\exists \underline{\nu}[\sigma(\underline{\mu}, \underline{\nu})] \wedge (\exists \underline{\nu}'[\sigma(\underline{\mu}', \underline{\nu}')] \rightarrow A(\underline{\mu}', x'))]]$, where A has no other free choice sequence variables, and $|\underline{\mu}|$ need not equal $|\underline{\nu}|$ and likewise for $\underline{\mu}'$ and $\underline{\nu}'$.

Proof:

Given any $\underline{\mu}$ assume that $\exists x[A(\underline{\mu}, x)]$.

By σ-μ-1 this implies that $\exists x \exists \sigma[\exists \underline{\nu}[\sigma(\underline{\mu}, \underline{\nu})] \wedge A'(\sigma, x)]$.

Which, by σ-μ-2 this implies that $\exists x \exists \sigma[\exists \underline{\nu}[\sigma(\underline{\mu}, \underline{\nu})] \wedge \forall \underline{\mu}'[\exists \underline{\nu}'[\sigma(\underline{\mu}', \underline{\nu}')] \rightarrow A(\underline{\mu}', x)]]$.

This is equivalent to $\exists \sigma \exists x' \forall \underline{\mu}'[\exists \underline{\nu}[\sigma(\underline{\mu}, \underline{\nu})] \wedge (\exists \underline{\nu}'[\sigma(\underline{\mu}', \underline{\nu}')] \rightarrow A(\underline{\mu}', x'))]$.

Discharging our assumption and quantification over $\underline{\mu}$ this gives us the following.

$\forall \underline{\mu}[\exists x[A(\underline{\mu}, x)] \rightarrow \exists \sigma \exists x' \forall \underline{\mu}'[\exists \underline{\nu}[\sigma(\underline{\mu}, \underline{\nu})] \rightarrow (\exists \underline{\nu}'[\sigma(\underline{\mu}', \underline{\nu}')] \rightarrow A(\underline{\mu}', x'))]]$ as required. ♠

Unlike in all the formal conventional theories discussed, our notion of weak continuity is derived straight from a more primitive notion – our axioms of knowledge – and these are **fully justified** based on the BHK interpretation of the quantifiers.

This completes the section on the links between knowledge states, choice sequences and predicates on choice sequences. We will now proceed to formally construct the language of knowledge states in the following sections and, from this, the universe of knowledge states.

§4.3. Atomic Knowledge States and Connectives

In the previous section, we outlined a theory of knowledge states in general without going into the specifics. In this section we shall remedy this and construct the formal syntax of knowledge states followed by their semantic interpretations. We will begin working only with knowledge states about a single choice sequence, and will generalise the idea to tuples of choice sequences later as needed.

Our approach will differ from the conventional theory where finite information is usually restricted to initial segments (extensional lists of terms). Instead, we will extend the notion of finite information to include intensional information, i.e. laws, spreads, fans and relationships with other choice sequences as well as initial segments. To do this, we shall define **seven** types of *atomic knowledge states* (two shall be covered in a later section) and a pair of connectives, so that we may construct any kind of finite information we desire by using connectives to combine atomic knowledge states.

The first kind of atomic information we need to be able to express is about a **specific element** of a sequence. As mentioned earlier, it does not matter where this type of information comes from, be it read from observing the generating process (our printed pile of elements) or 'hard coded' into it (the manual). Unlike the conventional theories mentioned previously, we do not differentiate between elements determined 'in advance' and elements generated 'afterwards'. This is why we insist that this type of atomic knowledge state covers both cases instead of just treating the 'hard coded' case as a special kind of spread.

If we had, for example, the information 'the x^{th} element is y' (note we are not mentioning a sequence at this point), we would require a standard way of expressing this information so that anyone reading it can always extract which element of the choice sequence we are talking about and what we are saying about it. We will use the syntax below to represent this.

$SE(x, y)$ (where SE abbreviates 'Set Element')

We also need to define the semantics for this kind of atomic knowledge; 'what restrictions on element generation does $SE(x, y)$ force'? Given that SE atoms are the **only non-empty knowledge states** we can generate by querying elements of a choice sequence, we must also ask 'what information must be given to us from querying a choice sequence to imply $SE(x, y)$ is consistent with that choice sequence'? The statement below answers both these questions and provides us with our semantics for such knowledge states.

$$\forall \mu[(SE(x, y))(\mu) \iff \mu(x) = y]$$

The second kind of atomic information we need to be able to express is about a **law** restricting a sequence. Our notion of a law is a constructive functions of type $N \mapsto N$; if you input an index, you obtain that element value. An example of the type of information we would like to express is 'for every x the x^{th} element is generated by the law f'. We require a standard way of expressing this information so that anyone reading it can always extract the law restricting the sequence. The syntax

$$Law(f)$$

performs this function perfectly.

We need to define the semantics for this kind of atomic knowledge. To explore this idea, we ask 'how would $Law(f)$ holding for a sequence restrict the elements of that sequence?'. If $Law(f)$ holds, then we can be certain that $\forall x \mu(x) = f(x)$; this can be written symbolically to give us our semantics as the following.

$$\forall \mu[(Law(f))(\mu) \rightarrow \forall x \mu(x) = f(x)]$$

This implication, as mentioned previously, is only **one way**. If a law extensionally agrees with the generating method of a choice sequence, we cannot imply that the choice sequence is intensionally generated by this law. Just because something looks like the law $\lambda x.2x$ does not mean that it is $\lambda x.2x$; it could be $\lambda x.\frac{8x}{4}$, for example.

The third kind of atomic information we need to be able to express is about a **spread restriction** on a choice sequence. Note that we use the conventional notion of a spread, i.e. spread laws act on finite sequences that form initial segments of choice sequences as defined in §1.2. Continuing on, the general case we wish to be able to represent is 'belongs to the spread defined by the spread law s'. We require a standard way of expressing this information so that anyone reading it can always extract the specific spread restricting the choice sequence. The syntax we will use is presented below.

$$Spread(s)$$

A constructive function acting on the knowledge state could parse the knowledge state and extract the spread law s.

We need to define the semantics for this kind of atomic knowledge state. Again, we ask the formative question, 'what restrictions on element generation would $Spread(s)$ impose?'. If $Spread(s)$ were intensional information for some μ, then we must have $\mu \in s$. Written symbolically this gives us our semantics as the following.

$$\forall \mu[(Spread(s))(\mu) \rightarrow \mu \in s]$$

The reader may wonder why we do not consider our, previously defined, *Law* atoms as a special case of *Spread* atom. In a way it is; every Law has an **extensionally equivalent** Spread atom. However, they are not intensionally equivalent as *Law* atoms contain an additional piece of information; the fact that there is only one path through this spread, i.e. there is a unique element for each index. There is also the fact that spreads, in general, need not be entirely constructive, whereas our laws must be by their very definition. Overall it seems better to keep our *Law* atoms; there is nothing to stop one using spreads should one desire, if one can tolerate the loss of information.

The fourth kind of atomic information we need to be able to express is about a special kind of spread restriction, namely a **fan restriction** on a sequence, for example 'belongs to the fan defined by the spread law s with constructive function f providing an upper limit on the number of children of each node'. We require a standard way of expressing this information so that anyone reading it can always extract the specific spread restricting the choice sequence as well as a function generating an **upper bound** on the element values possible at each node in the spread. The syntactic way to express this would be the following.

$$Fan(s, f)$$

A constructive function acting on the knowledge state could parse the knowledge state and extract the spread law s, and the function f capping the number of possible children each node can have.

To define the semantics of this type of atomic knowledge state, we must provide an answer to the question, 'what restrictions on element generation would $Fan(s, f)$ impose should it be consistent with a choice sequence?'. If $Fan(s, f)$ is consistent with a choice sequence μ, then we must have that both $\mu \in s$ and $\forall x[\mu(x) < f(\bar{\mu}(x))]$. Written symbolically, this gives us our semantics as the following.

$$\forall \mu[(Fan(s, f))(\mu) \rightarrow \forall x[\mu(x) < f(\bar{\mu}(x))] \wedge \mu \in s]$$

The fifth kind of atomic knowledge we wish to be able to express is an **absence of any kind of information** (we know 'nothing' about the object a sequence). Our reason for doing so is that such an absence of information is necessary as, in some cases, we will know literally 'nothing' about a choice sequence (a state rather akin to that which one holds initially for a [KT] proto-lawless sequence). We will require such a knowledge state for our notion of lawlessness (it essentially forms the 'maximal' knowledge state for our refined notion of lawless sequences which will be introduced in §4.6). This kind of 'information absence' cannot be constructed from our existing atoms, and thus requires its own syntax. The syntax $\sigma_\emptyset$ proves to be sufficient as it allows us to construct the empty knowledge state.

A constructive function acting on this knowledge state can quickly identify that it is the empty knowledge state. It is worth noting that an absence of knowledge is **syntactically distinct** from knowledge that is always true.

We also need to define the semantics of this kind of atomic knowledge. 'What kind of restrictions does no information place on the elements generated by a choice sequence?' is the question we require an answer to. In this instance it is clear that it imposes no restrictions and, in fact, what we observe is that the empty knowledge state is consistent with every choice sequence. This gives us our semantics as the following.

$$\forall \mu(\sigma_{\emptyset}(\mu))$$

Now, we have all of the atomic knowledge states pertaining to a single choice sequence, we require some connectives so that we may construct more complex knowledge states. We will use two connectives, $\sqcap$ and $\sqcup$, to do this.

Informally speaking, given two knowledge states σ_1 and σ_2 consistent with some choice sequence μ, $(\sigma_1 \sqcap \sigma_2)(\mu)$ iff both $\sigma_1(\mu)$ and $\sigma_2(\mu)$. More formally we write the following.

$$\sigma_1(\mu) \wedge \sigma_2(\mu) \iff (\sigma_1 \sqcap \sigma_2)(\mu)$$

Informally speaking, given two knowledge states σ_1 and σ_2 consistent with some choice sequence μ, $(\sigma_1 \vee \sigma_2)(\mu)$ is satisfied iff $\sigma_1(\mu)$ or $\sigma_2(\mu)$ are satisfied. More formally we write the following.

$$\sigma_1(\mu) \vee \sigma_2(\mu) \iff (\sigma_1 \sqcup \sigma_2)(\mu)$$

To ensure our formulae do not end up unbearably long we adopt the following notational conventions.

Given an arbitrary finite number of knowledge states $\sigma_1, \sigma_2, ..., \sigma_x$, we may represent the composite knowledge state $\sigma_1 \sqcap \sigma_2 \sqcap ... \sqcap \sigma_x$ by $\bigsqcap_{i=1}^{x} \sigma_i$.

Given an arbitrary finite number of knowledge states $\sigma_1, \sigma_2, ..., \sigma_x$, we may represent the composite knowledge state $\sigma_1 \sqcup \sigma_2 \sqcup ... \sqcup \sigma_x$ by $\bigsqcup_{i=1}^{x} \sigma_i$.

The reader may be wondering why we have no infinite cases for $\sqcap$ and $\sqcup$. We remind the reader that knowledge states represent **finite** information and are thus finite objects. Thus, if we were to have infinite cases for $\sqcap$ and $\sqcup$, we would not longer be dealing a finite object contradicting the entire notion.

Before bringing this section to a close we raise the issue of negation; why do we not consider objects such as $\neg SE(x, y)$ or $\neg Spr(s)$? The answer to this is, perhaps, slightly unsatisfying; we avoid negation as it has been proven to complicate matters with regards to conventional theories, and we do not wish to complicate a theory we are presenting as a viable notion that captures the conventional theory in full. However, it is worth noting that we can always obtain a spread extensionally equivalent to any negative piece of intensional information, so in this regard perhaps we do make some ground on the existing notion.

Finally, we introduce the following notational convention: **from this point onwards** σ_ϵ will be used to denote **non-empty** atomic knowledge states.

§4.4. Knowledge States About Tuples of Choice Sequences

Until now we have only introduced knowledge states about the generating process of a single choice sequence and, as seen from our earlier informal examples and numerous areas of the conventional theory, this is not sufficient. Many of the issues with the conventional theory of choice sequences arise when we attempt to relate choice sequences to one another, or deduce elements of one sequence from another using such a relation. Traditionally, statements about tuples of choice sequences are reduced down to statements about a single sequence formed by encoding the original choice sequences into a single choice sequence; we shall avoid this method as it reduces the clarity of ideas presented. It is with this in mind that we proceed with utmost caution and extend our language to multi-sequence knowledge states.

We begin by extending our syntax. Currently, if we have $(SE(0, 3))(\mu, \nu)$, we do not know if it is μ or ν that has its first element as 3. Earlier we mentioned that we will refer to sequences by their index based on the order they are mentioned in the knowledge state. Hence, μ would be our first sequence and ν would be our second sequence. Thus, we already have a framework in place to extend the syntax of our knowledge states by adding in an extra variable denoting the sequence the atom refers to. We can now clarify our earlier knowledge state to $(SE(1,0,3))(\mu,\nu)$ which tells us that, since ν is the second sequence, $\nu(0) = 3$.

In general our new syntax for atomic knowledge states is as follows.

$SE(w,x,y)$ which reads as 'the x^{th} element of choice sequence w is y'.

$Law(w,f)$ which reads as 'the constructive function f is intensionally defined to generate the elements of choice sequence w'.

$Spread(w,s)$ which reads as 'choice sequence w is intensionally defined as a member of the spread defined by spread law s'.

$Fan(w,s,f)$ which reads as 'choice sequence w is intensionally defined as a member of the fan defined by the spread law s with upper element limit determined by the constructive function f'.

The reader might note that the syntax for the empty knowledge state has not been altered. This is because we are saying we have no knowledge at all, not that we have no knowledge about a specific object.

As a notational convention, from now on, we will **omit** the sequence number in our syntax iff we are dealing with a knowledge state about **a single** choice sequence.

With this change to our syntax, our semantics also require some modification. The change is minor; we allow $\sigma(\mu)$ and our quantifiers, $\exists\mu$ and $\forall\mu$, to be extended to their tuple cases $\sigma(\underline{\mu})$, $\exists\underline{\mu}$ and $\forall\underline{\mu}$ respectively, where we remind the reader that $\underline{\mu}$ represents a finite tuple of choice sequences. Our semantics change as follows.

$$\forall\underline{\mu}[(SE(w,x,y))(\underline{\mu}) \iff \underline{\mu}_w(x) = y]$$

$$\forall\underline{\mu}[(Law(w,f))(\underline{\mu}) \to \forall x[\underline{\mu}_w(x) = f(x)]]$$

$$\forall\underline{\mu}[(Spread(w,s))(\mu) \to \underline{\mu}_w \in s]$$

$$\forall\underline{\mu}[(Fan(w,s,f))(\mu) \to \forall x[\underline{\mu}_w(x) < f(\bar{\underline{\mu}}_w(x))] \wedge \underline{\mu}_w \in s]$$

Some examples of these ideas in action are given below.

$$\forall\mu_0\forall\mu_1[(SE(0,2,3) \sqcap SE(1,3,4))(\mu_0,\mu_1) \iff \mu_1(2) = 3 \wedge \mu_2(3) = 4]$$

$$\forall\mu_0\forall\mu_1[(Law(0,\lambda x.2x) \sqcap Spread(1,s_{even}))(\mu_1,\mu_2) \to \forall x(\mu_0(x) = 2x) \wedge \mu_1 \in s_{even}]$$

$$\forall\mu_0\forall\mu_1[(SE(0,2,3) \sqcap Law(1,\lambda x.x))(\mu_0,\mu_1) \to \mu_0(2) = 3 \wedge \forall x(\mu_1(x) = x)]$$

We are now capable of forming knowledge states **about** different choice sequences. However, we still have no syntax to **relate** different choice sequences. We are now able to construct knowledge states such as $SE(0,1,1) \sqcap SE(1,1,1)$, and $Law(0,\lambda x.2x) \sqcap Law(1,\lambda x.2x-1)$. However, knowledge states such as 'the first elements of two sequences are the same', and 'the x^{th} element of the first sequence is always twice the x^{th} element of second sequence' are still beyond our grasp.

An intuitive way to approach the first of those two examples would to be use variables akin to those in a declarative programming language such as Prolog, i.e. something similar to $SE(0,0,X) \sqcap SE(1,0,X)$. This would allow us to equate two elements, but could we expect a continuous operation to parse this kind of knowledge in the way we mean it? A programming language such as Prolog would certainly be able to, via simple unification; so, it seems reasonable that a continuous operation would be able to do so as well.

Before proceeding to the second example, we define a very useful notational convention below.

We will write $\{w\}$ to denote the choice sequence with index w within a knowledge state.

The second example proves to be more troublesome and shows us that such variable unification would be infeasible for more complex examples. For the example, 'the x^{th} element of one sequence is always twice the x^{th} element of another', we would have $Law(0, f)$, where $f(x) = 2\{w_1\}(x)$. We would hit trouble with this method when we lack an element of w_1 in our knowledge state that is asked for by the law defining w_0. Our definition of a law requires us to generate any element on demand from the law and we would no longer be able to do so if we use this method.

Thus, with our options exhausted, it is clear that some alternative syntax is needed; a new notion to functionally relate a tuple of choice sequences must be defined. When we relate choice sequences, we must remember that any such relation must work with incomplete objects; thus, it seems reasonable to assert that any such relation must be expressible as a continuous function. As mentioned above, the strong notions of law and spread impose too many restrictions on the kind of relation possible; so, the format will be different from both. Given our simplest example, 'the first elements of two sequences are the same', if we write R for the relation $\{w_0\}(0) = \{w_1\}(0)$, then the syntax

$Rel(R, 0, 1)$, where Rel abbreviates 'Relation',

could be easily parsed by a continuous operation acting on the knowledge state, allowing it to identify which objects are being spoken about ($\{w_0\}, \{w_1\}$) and what is being said about them (the relation R holds). Continuing on to our second, previously more troublesome, example 'the x^{th} element of the first sequence is always twice the x^{th} element of the second sequence', we write

$Rel(R, 0, 1)$

where R is the relation $\forall x[\{w_0\}(x) = 2\{w_1\}(x)]$. Again, a continuous operation could parse this and extract the information required.

This, however, is not quite sufficient as we will now see. Adding a third example, 'for all x the sum of the x^{th} elements of two sequences is equal to the sum of the x^{th} elements of two other sequences' shows us a different kind of relation to the first two. In this form of relation, some algebraic manipulation is required to find any x^{th} element of a sequence using the x^{th} elements of the other two; this demands a lot from our notion of function! Thus we need to split our relations into two kinds, the first being relations that allow us to generate the elements of one sequence from information about a tuple of other sequences, we shall call these *Law Relations* as they behave in a similar way to Laws (in that they allow us to construct elements of the sequence). The second type of relation being relations between tuples of sequences that do not allow us to generate elements without some algebraic manipulation (if at all), we shall call these *Spread Relations* as they behave more closely to spreads. Hence our examples would now be written as follows.

> $LR(R, 0, 1)$, where R is the relation $\{0\}(0) = \{1\}(0)$, and LR abbreviates 'Law Relation'.

and

> $LR(R, 0, 1)$, where R is the relation $\forall x[\{0\}(x) = 2\{1\}(x)]$.

and

> $SR(R, 0, 1, 2, 3)$, where R is the relation $\forall x[\{0\}(x) + \{1\}(x) = \{2\}(x) + \{3\}(x)]$, and SR abbreviates 'Spread Relation'.

A point worth raising here is that there is a fundamental difference between the intensional information expressed by LR and SR; LR is indicating that you can generate the elements of one sequence via a tuple of other sequences, SR is merely indicating that there is some relationship between a tuple of sequences. For every LR atom $LR(R, w_0, w_1, ..., w_x)$ there is an extensionally equivalent SR atom $SR(R, w_0, w_1, ..., w_x)$; however the converse is not necessarily true (for example, the relation $SR(\forall x[\lfloor \frac{\{0\}(x)+\{1\}(x)}{2} \rfloor = \lfloor \frac{\{2\}(x)+\{3\}(x)}{2} \rfloor])$ most certainly cannot be expressed as a LR atom).

Our general syntax for both of these types of relationship are as follows.

> $LR(R, w_0, w_1, w_2, ..., w_x)$, where R is continuous operation (our generalised notion of continuity is covered in §4.7) mapping the finite tuple of sequences with indexes $w_1, w_2, ..., w_x$ to the sequence with index w_0.
>
> $SR(R, w_0, w_1, w_2, ..., w_x)$, where R is a continuous operation mapping the sequences with indexes $w_0, w_1, ...w_x$ to 0 (true) or 1 (false).

We also need to define the semantics for these kinds of atomic knowledge states. These are answers to the question 'what restrictions on element generation do these knowledge states impose'? If we assume our knowledge state is $LR(R, w_0, w_1, w_2, ..., w_x)$, where R maps $\underline{w}$ to w_i, our semantics would be the following.

$$\forall\underline{\mu}[(LR(R, w_0, w_1, w_2, ..., w_y))(\underline{\mu}) \rightarrow \forall x[\{w_0\}(x) = R(\{w_1\}, .., \{w_y\})]]$$

For the knowledge state $SR(R, \underline{w})$, where R is a spread relation between the tuple of sequences $\underline{w}$, the semantics would be the following.

$$\forall\underline{\mu}[(SR(R, w_0, w_1, ...w_y))(\underline{\mu}) \rightarrow R(\{w_0\}, \{w_1\}, ..., \{w_y\})]$$

We pause here to introduce a useful example, one that highlights the difference between LR and SR atoms

Example 4.4.1

> Define the relation R to be $R(\mu, \nu) \iff \forall x[\lfloor\frac{\mu(x)}{2}\rfloor = \lfloor\frac{\nu(x)}{2}\rfloor]$. This relation violates the requirements on R for it to be considered a law relation; since, no matter how much information we have about ν we will never be able to convert it into an element of μ.

This example shows us that not all spread relations are viable law relations (at least for our conception of law relation). However, it is worth noting that the converse (all law relations are also spread relations) does hold trivially.

Before continuing onto the next section, we will identify some knowledge states that may be problematic for continuous operations to handle. We will include them now to reassure the reader that we are aware of them and that we do have mechanisms in place to handle them (§6.4).

$\sigma \equiv LR(R, 0, 1) \sqcap LR(R', 1, 0)$, where R is the relation $\forall x[\{0\}(x) = 2\{1\}(x)]$ and R' is the relation $\forall x[\{1\}(x) = \frac{1}{2}\{0\}(x)]$. This knowledge state is interesting because we have a cycle, and if we want to know $\{0\}(0)$ then we need $\{1\}(0)$, to get $\{1\}(0)$ we need $\{0\}(0)$ and so on. The reader is advised to bear in mind that this type of knowledge state is not actually a problem because of the way we handle continuous operations on choice sequences.

$\sigma \equiv LR(R, 0, 1) \sqcap LR(R', 1, 0)$, where R is the relation $\forall x[\{0\}(x) = 2\{1\}(x)]$ and R' is the relation $\forall x[\{1\}(x) = 2\{0\}(x)]$. This knowledge state is only satisfiable by one pair of sequences (the case there $\{0\} = \{1\} =$ the sequence of $0s$); but would a function making use of this knowledge fail to terminate due to the inherent loop? The exposition in §6.4 indicates that such loops can be avoided.

$\sigma \equiv LR(R, 0, 1) \sqcap LR(R', 0, 1)$, where R is the relation $\forall x[\{0\}(x) = \{1\}(x) + 1]$ and R' is the relation $\forall x[\{1\}(x) = \{0\}(x) + 1]$. This knowledge state cannot be satisfied as it leads to $\{0\}(0) = \{1\}(0) + 1 = \{0\}(0) + 1 + 1$, and hence $\{0\}(0) = \{0\}(0) + 2$ which is contradictory. We can spot this contradiction fairly easily, however, would a function making use of this knowledge fail in some way because of it? Again, because of the way we handle continuous operations, the reader will see that this issue would simply not arise.

Finally, we introduce the following useful notational convention for a very useful Law Relation.

$LR(I, i, j)$ denotes the extensional equality relation between sequences with indexes i

and j (i.e. $(LR(I,0,1))(\mu,\nu) \to \forall x[\{0\}(x) = \{1\}(x)]$).

§4.5. Equality and Ordering of Knowledge States

In the conventional theory, where we restrict our notion of finite information to initial segments, our notions of ordering and equality are based on the lengths of initial segments and their elements. These notions give a partial ordering that is sufficient for the existing conventional theories. With our extended notion of finite information, the conventional notion is no longer sufficient and different notions of order and equality are required. Our finite information is dealt with in a strongly intensional manner and, as such, our equality and ordering must reflect this.

We informally define 'σ_1 is intensionally equal to σ_2' as 'σ_1 is syntactically identical to σ_2' and write this as $\sigma_1 \equiv \sigma_2$. We write $\sigma_1 \not\equiv \sigma_2$ to mean that $\sigma_1 \not\equiv \sigma_2$, that is, σ_1 and σ_2 are syntactically distinct. We read this as σ_1 *is not equivalent to* σ_2. So, we would consider $Law(\lambda x.x) \equiv Law(\lambda x.x)$ but $Law(\lambda x.x) \not\equiv Law(\lambda x.\frac{2x}{2})$ and $Law(\lambda x.x) \not\equiv Law(\lambda x.x) \wedge Law(\lambda x.x)$. Given any σ_1 and σ_2, we have that $\sigma_1 \equiv \sigma_2 \vee \sigma_1 \not\equiv \sigma_2$.

This notion of equality is very strong and of little use on its own. After we have defined our order axioms we will define a different, though related, notion of intensional equality that is weaker but far more useful.

Our notion for order of knowledge states is constructed by inductively defining the binary relation $\subseteq$ on knowledge states with the following axioms.:

O-1*a* $\forall\sigma_\epsilon[\sigma_\epsilon \subseteq \sigma_\epsilon]$

O-1*b* $\sigma_\emptyset \subseteq \sigma_\emptyset$

O-2*a* $\sigma_1 \subseteq \sigma_2 \to \sigma_1 \subseteq (\sigma_2 \sqcap \sigma_3)$

O-2*b* $\sigma_1 \subseteq \sigma_3 \to \sigma_1 \subseteq (\sigma_2 \sqcap \sigma_3)$

O-3 $(\sigma_1 \subseteq \sigma_3) \wedge (\sigma_2 \subseteq \sigma_3) \to (\sigma_1 \sqcap \sigma_2) \subseteq \sigma_3$

O-4a $\sigma_1 \subseteq \sigma_3 \rightarrow (\sigma_1 \sqcup \sigma_2) \subseteq \sigma_3$

O-4b $\sigma_2 \subseteq \sigma_3 \rightarrow (\sigma_1 \sqcup \sigma_2) \subseteq \sigma_3$

O-5 $(\sigma_1 \subseteq \sigma_2) \wedge (\sigma_1 \subseteq \sigma_3) \rightarrow \sigma_1 \subseteq (\sigma_2 \sqcup \sigma_3)$

O-6 $\subseteq$ is the minimum relation satisfying O-1 to O-5

We read $\sigma_1 \subseteq \sigma_2$ as 'σ_1 *is intensionally less than or equal to* σ_2'.

So given an arbitrary σ and σ', how do these axioms allow us to verify if $\sigma \subseteq \sigma'$? Let us look at a simple example

Example 4.5.1

Is $SE(0,1) \sqcup SE(1,2) \subseteq LAw(f) \sqcap SE(0,1)$?

Firstly we note that both our left hand knowledge state, and our right hand one, are comprised of finitely many atoms (two on each side) joined by finitely many connectives (one on each side). Also note that there is a relationship between the number of atoms and connectives; $the\ number\ of\ atomic\ knowledge\ states\ =\ the\ number\ of\ connectives\ +1$. This will always be the case because of how we construct our knowledge states.

Now, by O-2a and O-2b we know that it is sufficient for us to show that either $SE(0,1) \sqcup SE(1,2) \subseteq Law(f)$ (i) or $SE(0,1) \sqcup SE(1,2) \subseteq SE(0,1)$ (ii).

Note, we now have a choice between two simpler statements to prove, we have reduced the number of connectives on the right hand side to 0.

If we try to prove case (i) then, by O-4a and O-4b, it will be sufficient to prove that either $SE(0,1) \subseteq Law(f)$ (ia) or $SE(1,2) \subseteq Law(f)$ (ib).

Note that, again, we have reduced our number of connectives, this time on the left hand side, to 0.

We can no longer reduce the number of connectives on either side and thus the only viable axioms are O-$1a$ and O-$1b$, neither of which holds for either statement (ia) or (ib). So we can conclude that $SE(0,1) \sqcup SE(1,2) \not\subseteq Law(f)$

If we try to prove (ii), however, then by using a similar method we end up with the statement $SE(0,1) \subseteq SE(0,1)$ which **is** a case of O-$1a$ and thus $SE(0,1) \sqcup SE(1,2) \subseteq SE(0,1)$ and hence $SE(0,1) \sqcup SE(1,2) \subseteq LAw(f) \sqcap SE(0,1)$

In general we use the fact that, given two arbitrary knowledge states σ and σ', both will contain finitely many connectives. We then apply rules O-2-5 **finitely many times** to obtain cases where we have 0 connectives (look for a match to our statement on the right and then see what cases are sufficient to prove it by looking to the left, there will always be one or two) on both sides and use rules O-$1a$ and O-$1b$ to check for a 'pattern match'. Rules O-3 and O-5 give us two statements to verify rather than just one. We will always be able to tell if a relationship holds or not via finitely many applications of the rules O-1-5.

Theorem 4.5.1

$\forall\sigma[\sigma \subseteq \sigma]$

Proof:

All knowledge states are either atomic, or constructed via applying the connectives $\sqcap$ and $\sqcup$ to existing knowledge states. Thus to prove our assertion we first needs to show that it holds for atomic (and empty) knowledge states; we then need to show that our assertion also holds after finitely many applications of $\sqcap$ and $\sqcup$.

Base Cases : $\forall\sigma_\epsilon[\sigma_\epsilon \subseteq \sigma_\epsilon]$ and $\sigma_\emptyset \subseteq \sigma_\emptyset$ are satisfied by O-$1a$ and O-$1b$ respectively.

Induction Hypothesis (IH): $\sigma_1 \subseteq \sigma_1$ and $\sigma_2 \subseteq \sigma_2$

We now only need to prove, assuming our induction hypothesis, that $\sigma_1 \sqcap \sigma_2 \subseteq \sigma_1 \sqcap \sigma_2$ and $\sigma_1 \sqcup \sigma_2 \subseteq \sigma_1 \sqcup \sigma_2$.

Assuming (IH), $\sigma_1 \subseteq \sigma_1 \sqcap \sigma_2$ by O-2a and $\sigma_2 \subseteq \sigma_1 \sqcap \sigma_2$ by O-2b.

Hence, $\sigma_1 \sqcap \sigma_2 \subseteq \sigma_1 \sqcap \sigma_2$ by O-3 as required.

Likewise, we can show that $\sigma_1 \sqcup \sigma_2 \subseteq \sigma_1 \sqcup \sigma_2$.

Thus, by induction, $\forall\sigma[\sigma \subset \sigma]$, as required. ♠

Using $\subseteq$ we will define another type of relation between knowledge states, $\cong$. This relation is weaker than $\equiv$, but it is still strongly intensional and sufficient for our purposes. We present it formally below.

C-1 $\sigma_1 \cong \sigma_2 \iff \sigma_1 \subseteq \sigma_2 \wedge \sigma_2 \subseteq \sigma_1$

We write $\sigma_1 \subset \sigma_2$ to mean $(\sigma_1 \subseteq \sigma_2) \wedge (\sigma_1 \ncong \sigma_2)$. We read $\sigma_1 \subset \sigma_2$ as σ_1 *is intensionally less than* σ_2.

Finally, we comment that $\subseteq$ is decidable, as it involves direct syntax comparison and nothing else.

We now define a very useful piece of notation for atomic knowledge states that we will make extensive use of later on as follows.

$$\sigma_\epsilon \in \sigma \iff \begin{cases} \sigma \equiv \sigma_\epsilon \\ \quad or \\ \sigma \equiv \sigma' \sqcap \sigma'' \;\wedge\; (\sigma_\epsilon \in \sigma' \;\vee\; \sigma_\epsilon \in \sigma'') \\ \quad or \\ \sigma \equiv \sigma' \sqcup \sigma'' \;\wedge\; (\sigma_\epsilon \in \sigma' \;\vee\; \sigma_\epsilon \in \sigma'') \end{cases}$$

Theorem 4.5.2 : $\equiv \rightarrow \cong$

$\forall\sigma_1\forall\sigma_2(\sigma_1 \equiv \sigma_2 \rightarrow \sigma_1 \cong \sigma_2)$

Proof:

Trivial and left to the reader.

The converse of this theorem is false as proven by the counter-example $SE(2,3) \sqcap SE(1,2) \cong SE(1,2) \sqcap SE(2,3)$, but $SE(2,3) \sqcap SE(1,2) \not\equiv SE(1,2) \sqcap SE(2,3)$.

Theorem 4.5.3

The following results follow easily from our order axioms combined with C-1.

$L1$ $(\sigma_1 \sqcap \sigma_2) \sqcap \sigma_3 \cong \sigma_1 \sqcap (\sigma_2 \sqcap \sigma_3)$

$L2$ $(\sigma_1 \sqcup \sigma_2) \sqcup \sigma_3 \cong \sigma_1 \sqcup (\sigma_2 \sqcup \sigma_3)$

$L3$ $\sigma_1 \sqcap \sigma_2 \cong \sigma_2 \sqcap \sigma_1$

$L4$ $\sigma_1 \sqcup \sigma_2 \cong \sigma_2 \sqcup \sigma_1$

$L5$ $\sigma_1 \sqcap (\sigma_1 \sqcup \sigma_2) \cong \sigma_1$

$L6$ $\sigma_1 \sqcup (\sigma_1 \sqcap \sigma_2) \cong \sigma_1$

Proof:

We only provide a proof of $L6$ to illustrate the general method, and all of these proofs can be considered trivial.

$\sigma_1 \preceq \sigma_1$ by O-1 and hence $\sigma_1 \subseteq \sigma_1 \sqcap \sigma_2$ by O-2a.

Thus we have $\sigma_1 \subseteq \sigma_1 \sqcup (\sigma_1 \sqcap \sigma_2)$ by O-5 and $\sigma_1 \sqcup (\sigma_1 \sqcap \sigma_2)$ by O-4a.

Hence by C-1 we obtain $\sigma_1 \sqcup (\sigma_1 \sqcap \sigma_2) \cong \sigma_1$ as required. ♠

These results imply that intensional ordering on knowledge states forms a lattice, and hence a partial ordering. This result is expected as we have no way to 'measure' knowledge, and a total ordering would imply that any two pieces of information are directly related in the ordering. How would one relate two *Law* atoms, or a *Law* atom to a *SE* atom? The fact that our ordering is partial has no detrimental impact on the theory and makes sense, as not all types of knowledge are directly comparable.

Finally, we introduce two easily provable results dealing with $\sigma_\emptyset$ below.

$$\forall\sigma[\sigma \sqcap \sigma_\emptyset \cong \sigma_\emptyset \sqcap \sigma \cong \sigma]$$

'If we know nothing and something then we know something'.

$$\forall\sigma[\sigma \sqcup \sigma_\emptyset \cong \sigma_\emptyset \sqcup \sigma \cong \sigma_\emptyset]$$

'If we know nothing or something then we know nothing'.

§4.6. Special Species of Knowledge States, OMR and KS-lawlessness

In this section, we will consider the matter of contradictory knowledge states and their impact on density. Following this, we will formally define two useful species of knowledge states that we will use later on (Σ_{SE} and Σ_{IS}), the one manual rule (OMR), and our notion of lawlessness (KS-lawlessness). These matters may all seem unrelated but the reader is assured that there is a strong common ground for these notions.

As we mentioned earlier, we do not ban '*inconsistent*' knowledge states (knowledge states that describe intensional information about generating processes that are impossible to carry out), as we cannot always identify them. It is not always decidable if a given knowledge state is consistent with some given choice sequence or if it describes an impossible process. For example, given the spread of sequences of perfect numbers (s_{perf}) and the spread of sequences of odd numbers (s_{odd}) then $Spread(s_{perf}) \sqcap Spread(s_{odd})$ would only be consistent with the generating process of a choice sequence comprising odd perfect numbers, but we do not know if such odd perfect numbers exist. Both of these spreads are consistent with a species of individual choice sequences; finding one that satisfies their combination is uncertain at best.

The presence (and inability to identify) such inconsistent knowledge states denies us the density statement $\forall\sigma\exists\mu[\sigma(\mu)]$ and seemingly presents a major problem for our notion of a choice sequence. I would argue otherwise; the notion of knowledge density (for all knowledge there is some choice sequence consistent with it) is far stronger than the Baire density (for all finite sequences there is some choice sequence containing it as an initial segment); it will be shown later in this section that we can argue strongly for Baire density even with our conception of a choice sequence and later, in chapter 5, that full knowledge density is simply not required to provide a foundation for analysis. To take the first step towards this, we will define the species of knowledge states that is the direct equivalent to the species of initial segments in the conventional theory.

First, we define inductively Σ_{SE_0}, the species of knowledge states comprising only finite conjunctions of SE atoms, which are as follows.

Σ_{SE_0}-1 $\sigma_\emptyset \in \Sigma_{SE_0} \wedge \forall w \forall x \forall y[SE(w,x,y) \in \Sigma_{SE_0}]$

Σ_{SE_0}-2 $\forall \sigma \in \Sigma_{SE_0} \forall \sigma' \in \Sigma_{SE_0}[\sigma \sqcap \sigma' \in \Sigma_{SE_0}]$

Σ_{SE_0}-3 Σ_{SE_0} is the minimum species of knowledge states satisfying Σ_{SE_0}-1 and Σ_{SE_0}-2.

Because of how elements of Σ_{SE_0} are constructed we define the species of contradictory Σ_{SE_0} knowledge states as follows.

$$\forall \sigma \in \Sigma_{SE_0}[\sigma \in \Sigma_\perp \iff \exists w \exists x \exists y \exists z \neq y[SE(w,x,y) \wedge SE(w,x,z) \subseteq \sigma]]$$

The statement above is clearly decidable; all we need to do is compare each SE atom with every other SE atom. Thus, we can conclude that for Σ_{SE_0} (and hence any subspecies of Σ_{SE_0}) inconsistency is decidable. We present this formally below.

$$\forall \sigma \in \Sigma_{SE_0}[\sigma \in \Sigma_\perp \vee \sigma \notin \Sigma_\perp]$$

From this we can thus construct our first useful species of knowledge states Σ_{SE}, the set of consistent knowledge states comprising only finite conjunctions of SE atoms. We present this formally below.

Σ_{SE}-1 $\sigma \in \Sigma_{SE} \iff \sigma \in \Sigma_{SE_0} \wedge \sigma \notin \Sigma_\perp$

Σ_{SE}-2 Σ_{SE} is the minimum species of knowledge states satisfying Σ_{SE}-1.

Using our definition of Σ_{SE}, we can construct our second useful species of knowledge states Σ_{IS}, the set of knowledge states comprising only finite conjunctions of SE atoms that represent initial segments of choice sequences (finite sequences). We define this formally below.

Σ_{IS}-1 $\forall\sigma[\sigma \in \Sigma_{SE} \wedge \forall x > 0\ \exists w[\exists y[SE(w,x,y)] \in \sigma \rightarrow \exists y_0[SE(w,x-1,y_0) \in \sigma]]$

$\iff \sigma \in \Sigma_{IS}]$

Σ_{IS}-2 Σ_{IS} is the minimum species of knowledge states satisfying Σ_{IS}-1.

We define a useful function, MEI (Maximum Element Index), below.

$$MEI(\sigma) = \begin{cases} max(x \mid SE(w,x,y) \in \sigma)\ if\ \sigma \in \Sigma_{SE} \\ \nabla\ otherwise \end{cases}$$

Two results that are of interest are given below.

Theorem 4.6.1

$\forall\mu\forall\sigma \in \Sigma_{SE}[\sigma(\mu) \rightarrow \exists\sigma' \in \Sigma_{IS}[\sigma'(\mu) \wedge \sigma \subseteq \sigma']]$

'Given any list of elements of a sequence we can fill in the gaps of that sequence to construct an initial segment of that sequence'.

Proof:

Trivial and left to the reader.

Theorem 4.6.2

$\forall\sigma \in \Sigma_{SE}[|\sigma| = 1 \rightarrow \exists\mu[\sigma(\mu)]]$ (density)

Proof:

Given any $\sigma \in \Sigma_{SE}$ such that $|\sigma| = 1$,

$\sigma \equiv SE(0,x_0,y_0) \sqcap SE(0,x_1,y_1) \sqcap ... \sqcap SE(0,x_z,y_z)$ by the definition of $\sigma \in \Sigma_{SE}$ and the fact that $|\sigma| = 1$.

By the fact that $\sigma \notin \Sigma_\perp$ we know that $\forall i \leq x \forall j \leq z[x_i = x_j \iff y_i = y_j]$.

Therefore we can define μ such that $\mu(x_0) = y_0 \wedge \mu(x_1) = y_1 \wedge ... \wedge \mu(x_z) = y_z$, and be certain that we are not forcing a single element to adopt two different values.

σ records the generating process of μ, hence $\sigma(\mu)$ by the definition of consistency.

Discharging our assumption and quantification over σ we obtain

$\forall \sigma \in \Sigma_{SE}[|\sigma| = 1 \rightarrow \exists \mu[\sigma(\mu)]]$ as required. ♠

A useful property of Σ_{IS} is as follows.

$\forall \sigma \in \Sigma_{IS} \exists \underline{n} \forall \underline{\mu}[\sigma(\underline{\mu}) \iff \underline{n} \subset \underline{\mu}]$ where $\underline{n} \subset \underline{\mu}$ abbreviates $\underline{n}_0 \subset \underline{\mu}_0 \wedge ... \wedge \underline{n}_{|\underline{\nu}|} \subset \underline{\mu}_{|\underline{\nu}|}$

For the finite tuple of sequences $\underline{n}$, and some $\sigma \in \Sigma_{IS}$, we write $\sigma \sim \underline{n}$ to mean $\forall w \forall x \forall y[SE(w, x, y) \in \sigma \iff x \leq len(n_w) \wedge w \leq |\underline{n}| \wedge n_w(x) = y]$, where $|\underline{n}|$ is the number of sequences in the tuple $\underline{n}$. In other words the knowledge state σ describes exactly the same information as the finite sequences $\underline{n}$. We will write $\sigma \sim n$ when $|n| = 1$.

Theorem 4.6.3

$\forall \sigma \in \Sigma_{IS}[|\sigma| = 1 \rightarrow \exists n[\sigma \sim n]]$

Proof:

Given any $\sigma \in \Sigma_{IS}$ assume $|\sigma| = 1$.

$\sigma \equiv SE(0, 0, x_0) \wedge SE(0, 1, x_1) \sqcap ... \sqcap SE(0, 2, x_z)$ by the definition of $\sigma \in \Sigma_{IS}$ and the fact that $|\sigma| = 1$.

Define the sequence $n = \langle x_0, x_1, ..., x_z \rangle$.

$\forall x \forall y[SE(x, y) \in \sigma \iff x \leq len(n) \wedge n(x) = y]$, which is sufficient for $n \sim \sigma$.

Discharging our assumption and quantification over σ we obtain

$\forall n \exists \sigma \in \Sigma_{IS}[|\sigma| = 1 \wedge \sigma \sim n]$ as required. ♠

Theorem 4.6.4

$\forall n \exists \sigma \in \Sigma_{IS}[|\sigma| = 1 \wedge \sigma \sim n]$

Proof:

Given any n such that $len(n) = i$,

Define $\sigma \equiv SE(0, 0, n(0)) \sqcap SE(0, 1, n(1)) \sqcap ... \sqcap SE(0, i, n(i))$.

$\sigma \in \Sigma_{SE}$ by definition and $|\sigma| = 1$.

$\forall x \forall y[SE(x, y) \in \sigma \iff x \leq i \wedge n(x) = y]$, which is sufficient for $n \sim \sigma$.

Discharging our assumption and quantification over σ we obtain

$\forall n \exists \sigma \in \Sigma_{IS}[|\sigma| = 1 \wedge \sigma \sim n]$ as required. ♠

Lemma 4.6.1

$\forall \sigma \in \Sigma_{SE} \forall \mu \forall \nu[\mu = \nu \iff (\sigma(\mu) \iff \sigma(\nu))]$

Proof:

Follows trivially by the definition of $\mu = \nu$. ♠

Lemma 4.6.2

$\forall n \forall m \forall \sigma \in \Sigma_{IS} \forall \sigma' \in \Sigma_{IS}[n \leq m \wedge \sigma \sim n \wedge \sigma' \sim m \rightarrow \sigma \subseteq \sigma']$

Proof:

Follows trivially by the definition of $\sigma \sim n$. ♠

As the reader may have already guessed, Σ_{IS} forms a good model for finite information about *KT-LL* and *M-LL* sequences in the conventional theory. However, at the moment, we lack a good model for the finite information about *GC* sequences, and it is for this task we now inductively define the following universe of knowledge states, Σ_{GC}. Please note, we follow

Van Der Hoeven (1982) by only allowing one sequence to be generated by up to two others; this is done purely for readability, this notion could easily be generalised to allow any finite tuple of sequences to generate another.

Σ_{GC}-1: $\forall\sigma[\sigma \in \Sigma_{IS} \rightarrow \sigma \in \Sigma_{GC}]$

"Initial segments are valid finite information for GC sequences"

Σ_{GC}-2: $\forall w \forall x \forall y \forall R_{\forall y[R|_y \in K}[LR(R, w, x) \in \Sigma_{GC} \wedge LR(R, w, x, y) \in \Sigma_{GC}]$

"The fact that a K-function generates the elements of one sequence from another (or two others) is valid finite information for GC sequences". We note here that the arity of R does not matter as we can simply interlace our sequences.

Σ_{GC}-3:

$$\begin{aligned}
\forall\sigma \in \Sigma_{GC} \forall R_{\forall i[R|_i \in K]} \forall w \forall x \forall y[\neg(\exists x' \exists y' \exists R'_{\forall i[R'|_i \in K]}[\\
& LR(R', w, x') \in \sigma \vee LR(R', w, x', y') \in \sigma \\
& \vee LR(R', x, x') \in \sigma \vee LR(R', x, x', y') \in \sigma \\
& \vee LR(R', y, x') \in \sigma \vee LR(R', y, x', y') \in \sigma \\
& \vee SE(x, x', y') \in \sigma \vee SE(y, x', y') \in \sigma] \\
& \rightarrow ((\sigma \sqcap LR(R, w, x)) \in \Sigma_{GC} \wedge (\sigma \sqcap LR(R, w, x, y)) \in \Sigma_{GC}))]
\end{aligned}$$

"The binding of a sequence to 'fresh' (no previous relations or elements defined) sequences is valid finite information for GC sequences"

Σ_{GC}-4:

$$\begin{aligned}
\forall\sigma \in \Sigma_{GC} \forall w \forall \forall y[(\neg(\exists x' \exists y' \exists R_{\forall i[R|_i \in K]}[LR(R, w, x') \in \sigma \vee LR(R, w, x', y') \in \sigma]) \\
& \wedge \forall x' < x \exists y[SE(w, x, y) \in \Sigma]) \\
& \rightarrow (\sigma \sqcap SE(w, x, y)) \in \Sigma_{GC}]
\end{aligned}$$

"Elements generated for any sequence that isn't determined by another is valid finite information for GC sequences"

Σ_{GC}-5 : Σ_{GC} is the minimum species of knowledge states satisfying Σ_{GC}-1-4

The species of knowledge stages Σ_{GC} grants us access to all viable finite information about GC sequences; thus we can define the universe of GC sequences (M_{GC}) to be $M_{GC} \equiv \{\underline{\mu} \mid \forall\sigma[\sigma(\underline{\mu}) \rightarrow \sigma \in \Sigma_{GC}]\}$. We note here just how restrictive the universe of GC sequences is; it is a far cry from the much more general notion of choice sequence we make use of. The absence of any *Spread*, *Fan* and *Law* atoms is not as noteworthy as one would hope as the first and second are tacitly present when we relativise our universe of choice sequences for analysis and the third is defined as a separate species of sequences in the same theory (the "lawlike" sequences in CS). What is of note, is the restrictions imposed on the kind of relations $R_{\forall i[R|_i \in K]}$, the absence of SR atoms, and the fact that the order in which we are given information is deemed to matter; these provide stark differences to our universe of choice sequences and that of GC sequences.

As promised in §4.2.1 we are now able to give a formal definition of the 'one manual rule' schema that will be used to give us a sub-universe of choice sequences suitable for analysis in §6.1.

OMR $\forall\underline{\mu}\exists\sigma[\forall\sigma' \subset \sigma[\sigma' \notin \Sigma_{SE}] \wedge \sigma(\underline{\mu}) \wedge \forall\sigma'[\forall\sigma'' \subseteq \sigma'[\sigma'' \notin \Sigma_{SE}] \wedge \sigma'(\underline{\mu}) \rightarrow \sigma' \subset \sigma]]$

(Notice that the universe of KS-LL sequences tacitly assume OMR. For a KT-LL choice sequence we know that the maximal knowledge state will be $\sigma_\emptyset$. When dealing with GC sequences we are always treating our sequences as if they were lawless, at least when considered alone. Thus one could argue for some assumption of OMR when dealing with GC sequences alone.)

We will now proceed to formally define our notion of lawlessness; that of a KS-lawless choice sequence. A notion that, I believe, is far more reasonable than any other in existence at the time of writing.

A choice sequence μ is *KS-lawless* iff $\forall\sigma[\sigma(\mu) \to \sigma \in \Sigma_{SE}]$.

Note that this still allows knowledge states for a KS-lawless $\mu \in M$ and some other $\nu \in M$ such as $(LR(R, 0, 1) \sqcap Law(1, f))(\mu, \nu)$; however, it does forbid cases such as $(LR(R, 0, 1) \sqcap Spr(0, s))(\mu, \nu)$. We shall denote the universe of KS-lawless sequences by M_{KSLS}, i.e. $\mu \in M_{KSLS} \iff \forall\sigma[\sigma(\mu) \to \sigma \in \Sigma_{SE}]$

Before we proceed to the next section, we introduce a very important claim below.

Claim

$\exists KS\text{-}LS$ $\forall\mu\exists\nu \in M_{KSLS}[\mu = \nu]$

Argument:

Given any μ, define a ν such that $\forall x[\nu(x) = \mu(x)]$; that is a ν such that $LR(I, 0, 1)(\mu, \nu)$ holds, where I is the identity law relation.

Now, currently, we possess no additional intensional information about ν; indeed, we do not know if there is any more. Under OMR, we are entitled to say that each choice sequence has a maximal knowledge state; in effect, we can place a second order restriction on a knowledge state preventing any additional non-element information about it (alone) being present. Thus, if we impose this restriction on ν, it would have a maximal knowledge state of $\sigma_\emptyset$ and the following would then hold: $\forall\sigma[\sigma(\nu) \to \sigma \in \Sigma_{SE}]$, i.e. $\nu \in M_{KSLS}$.

This argument definitely indicates $\exists KS\text{-}LS$ is plausible, however it falls short of actually providing a proof via OMR. As we have no direct proof of this concept yet (though it is conceivable one may exist), we will take it as an axiom when we define analysis. While it may be tempting to leave out OMR and simply include $\exists KS\text{-}LS$ in a formal theory, this would lose us the additional benefits of OMR outlined in §4.2.1.

§4.7. Extended Neighbourhood Functions – $\hat{K}_0$

In the conventional theory, we use the notion of B-continuity outlined in §1.4. For our theory, as we now consider laws, spreads, fans and relations with other choice sequences as usable finite information, we need a more general notion of continuity that expresses a new idea of choice sequence 'closeness' in terms of sharing some knowledge. We shall call this generalised notion of continuity Σ-*continuity* and define it below.

> ψ is Σ-*continuous* iff $\forall \underline{\mu}_{|\underline{\mu}|=|\psi|} \exists \sigma [\exists \underline{\nu} [\sigma(\underline{\mu}, \underline{\nu})] \wedge \forall \underline{\mu}' [\exists \underline{\nu}' [\sigma(\underline{\mu}', \underline{\nu}')] \rightarrow \psi(\underline{\mu}') = \psi(\underline{\mu})]]$, where $|\psi|$ is the number of choice sequence variables in ψ.

It may be unclear why, unlike the conventional theory, we allow the additional $\exists \underline{\nu}$ quantifiers on the RHS. The reason behind this is that their exclusion relies on the assumption that the evaluation of ψ does not rely on other choice sequences. This assumption lies at the heart of many results in the theory (as seen in §2.4) and given that we desire our system to be as general as possible, it makes better sense to develop the theory without making such assumptions.

Because our notion of continuity is different, we need a new way of 'representing' these continuous operations since, as seen above, when applying these Σ-continuous operations to choice sequences what really matters is the **finite information** we have about the choice sequences in question. We must consider the cases where we have **too little information** to evaluate the continuous operation or we have **conflicting information** that offers more than one possible solution (a new circumstance we must consider in the extended theory). Our solution to these dilemmas is found by using an *extended neighbourhood function* to represent a continuous operation as done in the conventional theory.

Within the extended theory, we informally define an extended neighbourhood function representing a continuous operation as some constructive function that outputs a natural number if the information supplied is sufficient to perform the continuous operation, ∇ if the finite information is **insufficient** to perform the continuous operation and Δ if the information supplied is **contradictory**.

In the extended theory our notion of 'finite information' is extended from 'initial segments' to **knowledge states** constructed from our atoms and connectives defined earlier. A quick note that we denote the **minimum** number of choice sequences contained in a knowledge state required to obtain a natural number output when $\hat{e}$ is applied by $ari(\hat{e})$. An *extended neighbourhood function* is a function $\hat{e}$ of type $\Sigma \mapsto N^*$ (for the definition of N^* please see §2.2) that satisfies the following conditions.

$\hat{K}_0 1$ $\forall \underline{\mu}_{|\underline{\mu}| \geq ari(\hat{e})} \exists \sigma [\exists \underline{\nu} [\sigma(\underline{\mu}, \underline{\nu})] \wedge \hat{e}(\sigma) \in N]$ (totality)

$\hat{K}_0 2$ $\forall \sigma \forall \sigma' [\sigma \subseteq \sigma' \rightarrow \hat{e}(\sigma) \preceq \hat{e}(\sigma')]$ (monotonicity)

$\hat{K}_0 3$ $\forall \sigma [\exists \underline{\mu} [\sigma(\underline{\mu})] \rightarrow \hat{e}(\sigma) \prec \Delta]$ (contradictoriness)

Any function satisfying these criteria is a member of the species of extended neighbourhood functions $\hat{K}_0$. This species is an extension of the one we constructed for the conventional theory as it handles more types of finite information and contradictory knowledge states. $\hat{K}_0 2$ is an extension of $K_0 2$, as the extension of a knowledge state that gives a natural number output may give a contradiction if the extension is contradictory. $\hat{K}_0 3$ places a strong restriction on the contradictory output, forbidding it for all knowledge states consistent with a choice sequence. This, in effect, means that we may **not** output Δ unless we are certain of, and can specifically indicate, a contradiction. The reader may wonder why we do not enforce $|\sigma| = |\underline{\mu}|$ in $\hat{K}_0 3$ and this is simply because,

$|\underline{\mu}| > |\sigma|$ would still indicate a $|\sigma|$-tuple of sequences with which σ is consistent (remember, we simply assume that we have no knowledge of the other sequences if $|\underline{\mu}| > |\sigma|$).

$|\underline{\mu}| < |\sigma|$ is excluded via our rules of consistency formation.

We formally define an extended neighbourhood function $\hat{e}$ (of type $\Sigma \mapsto N^*$) as *representing some continuous operation* ψ (of type $M^i \mapsto N$) iff $\forall\underline{\mu}\exists\sigma[\exists\underline{\nu}[\sigma(\underline{\mu},\underline{\nu})] \wedge \hat{e}(\sigma) = \psi(\underline{\mu})]$. We write this relation as $\psi \sim \hat{e} \iff \forall\underline{\mu}\exists\sigma[\exists\underline{\nu}[\sigma(\underline{\mu},\underline{\nu})] \wedge \hat{e}(\sigma) = \psi(\underline{\mu})]$, read as '$\psi$ is represented by $\hat{e}$'.

The additional $\exists\underline{\nu}$ in $\hat{K}_0 1$ is a direct consequence of our definition of Σ-continuity. This drastically widens the definition of what is, and what is not, construed as a $\hat{K}_0$ function; though, to what degree this impacts the development of certain ideas in the theory we cannot be certain at this time.

We will conclude this section with a theorem below which shows that $\hat{K}_0 3$ has an equivalent (and more usable) form.

$\hat{K}_0 3$ Alt $\forall\sigma\forall\sigma'[\exists\underline{\mu}[\sigma(\underline{\mu}) \wedge \sigma'(\underline{\mu})] \rightarrow \exists x \in N[\hat{e}(\sigma) \preceq x \wedge \hat{e}(\sigma') \preceq x]]$

Theorem 4.7.1

Assuming $\hat{e}$ satisfies $\hat{K}_0 2$, then $\hat{e}$ satisfies $\hat{K}_0 3 \iff \hat{e}$ satisfies $\hat{K}_0 3$ Alt.

Proof:

Assume $\hat{e}$ satisfies $\hat{K}_0 2$ and $\hat{K}_0 3$. Assume that given some arbitrary σ and σ', there exists some $\underline{\mu}$ such that $(\sigma \sqcap \sigma')(\underline{\mu})$.

By $\hat{K}_0 3$ this implies that $\hat{e}(\sigma \sqcap \sigma') \prec \Delta$.

This may then be rewritten as $\exists x[\hat{e}(\sigma \sqcap \sigma') \preceq x]$ because $A \prec \Delta \rightarrow \exists x[A \preceq x]$.

$\sigma \subseteq (\sigma \sqcap \sigma')$ by definition and hence by $\hat{K}_0 2$ we have that $\hat{e}(\sigma) \preceq \hat{e}(\sigma \sqcap \sigma') \preceq x$. This may be shown similarly for σ'.

Thus, we have that $\exists x(\hat{e}(\sigma) \preceq x \wedge \hat{e}(\sigma') \preceq x)$, and hence by discharging our initial assumption and our quantifiers $\forall\sigma\forall\sigma'[\exists\underline{\mu}[\sigma(\underline{\mu}) \wedge \sigma'(\underline{\mu})] \rightarrow \exists x[\hat{e}(\sigma) \preceq x \wedge \hat{e}(\sigma') \preceq x]]$ as required.

The converse is much simpler. Assume $\hat{K}_0 3$ Alt is true and set $\sigma \equiv \sigma'$, then we have $\forall\sigma[\exists\underline{\mu}[\sigma(\underline{\mu})] \rightarrow \exists x[\hat{e}(\sigma) \preceq x]$. Because $x \prec \Delta$ this may be re-written as $\forall\sigma(\exists\underline{\mu}\sigma(\underline{\mu}) \rightarrow \hat{e}(\sigma) \prec \Delta)$ which is our required converse.

Hence $\hat{K}_0 3 \iff \hat{K}_0 3$ Alt as required. ♠

§4.8. A Stronger Axiom of Continuity

In the conventional theory we rely on a stronger continuity axiom than open data, namely *BC-N*. To obtain the full strength of analysis, we will have to formulate (and justify) an analogue of this in our new language for our new notion of continuity.

Assume that $\forall\underline{\mu}\exists x[A(\underline{\mu}, x)]$ holds and A has no other choice sequence variables. The quantifier combination $\forall\underline{\mu}\exists x$ is read 'given any $\underline{\mu}$ we can **construct** an x', indicating that there is some constructive method for obtaining x from $\underline{\mu}$. Now if we, as in the conventional theory, assume that our continuous operations are all constructive (a reasonable assumption to make) we have some continuous operation ψ such that $\psi(\underline{\mu}) = x$. Such a method would have to rely only on finite information about $\underline{\mu}$, i.e. some knowledge state consistent with $\underline{\mu}$, and hence we make use of some extended neighbourhood function $\hat{e}$ such that $\psi \sim \hat{e}$. We can write this formally as follows.

$$\exists\hat{e} \in \hat{K}_0\forall\underline{\mu}\exists\sigma[\sigma(\underline{\mu}) \wedge A(\underline{\mu}, \hat{e}(\sigma))]$$

A remaining qualm is the case where x is derived in some way using information from some additional choice sequence. Conventionally (as we saw in §2.4) such additional information is considered not to exist due to the various arguments laid out relying on the extensionality of the predicates in question. This form of argument is not valid here, as our predicates are not assumed extensional and thus an alternative reformulation of our conclusion is required, and is given below.

$$\exists \hat{e} \in \hat{K}_0 \forall \underline{\mu} \exists \sigma [\exists \underline{\nu} [\sigma(\underline{\mu}, \underline{\nu})] \wedge A(\underline{\mu}, \hat{e}(\sigma)]$$

This formulation has the advantage of being fully justified by the meanings of the quantifiers, though the drawback is that this concept is weaker. How much weaker remains an open question, though the exposition on the derivation of $AC\text{-}NN$ given later indicates that this concept is, alone, not as strong as we may desire. For now, we give our fully justified axiom of continuity below.

$BC\text{-}N\text{-}KS$ $\forall \underline{\mu} \exists \underline{x} [A(\underline{\mu}, \underline{x})] \rightarrow \exists \underline{\hat{e}} \in \hat{K}_0 \forall \underline{\mu} \exists \sigma [\exists \underline{\nu} [\sigma(\underline{\mu}, \underline{\nu})] \wedge A(\underline{\mu}, \underline{\hat{e}}(\sigma))]$

where A has no other free choice sequence variables.

One lamentable side effect of the additional choice sequences ($\underline{\nu}$) is that we can no longer prove $AC\text{-}NN$, the argument going as follows:

Assume $BC\text{-}N\text{-}KS^+$ and the hypothesis of $AC\text{-}NN$ ($\forall x \exists y [A(x, y)]$).

Define $\mu \in M_N \iff \exists x [\mu \equiv \lambda z.x]$ (the species of sequences defined by a specific form of constant function).

Given any x define $\mu_x \equiv \lambda z.x$. Clearly, every $\mu \in M_N$ can be constructed in this way.

Define the predicate A' as $A'(\mu, y) \iff A(\mu(0), y)$.

Thus, using these definitions, we can rewrite the hypothesis of $AC\text{-}NN$ as $\forall \mu \in M_N \exists y [A'(\mu, y)]$.

Hence, by $BC\text{-}N\text{-}KS^+$ (suitably relativised to the spread M_N), $\exists \hat{e} \in \hat{K}_0$ such that $\forall \mu \in M_N \exists \sigma [\exists \underline{\nu} \in M_N [\sigma(\mu, \underline{\nu})] \wedge A'(\mu, \hat{e}(\sigma))]$ (i).

Because we know that $\forall\mu \in M_N \exists x[\mu \equiv \mu_x]$, that $A'(\mu, y) \iff A(\mu(0), y)$, the latter part of (i) can be rewritten as $\forall x \exists \sigma[\exists \underline{\nu} \in M_N[\sigma(\mu_x, \underline{\nu})] \wedge A(x, \hat{e}(\sigma))]$.

Now we can be certain that σ relies solely on x; all that remains is to produce an algorithm for deriving σ from x.

It is here that the proof breaks down. The $\exists \underline{\nu} \in M_N[\sigma(\mu_x, \underline{\nu})]$ part of the formulae really breaks down into $\exists z_0 \exists z_1 ... \exists z_i[\sigma(\mu_x, \nu_{z_0}, \nu_{z_1}, ..., \nu_{z_i})]$. For simplicity, assume that we only have one additional sequence and hence this is $\exists z[\sigma(\mu_x, \nu_z)]$. We know that all information about ν_z (when considered alone) is going to rely on z; so if we find z, then we can find our information.

So now we have the statement $\forall x \exists \sigma[\exists z[\sigma(\mu_x, \nu_z)] \wedge A(x, \hat{e}(\sigma))]$ which is equivalent to $\forall x \exists \sigma \exists z[\sigma(\mu_x, \nu_z) \wedge A(x, \hat{e}(\sigma))]$ which, in turn, is equivalent to $\forall x \exists z \exists \sigma[\sigma(\mu_x, \nu_z) \wedge A(x, \hat{e}(\sigma))]$ which implies $\forall x \exists z \exists \sigma[\sigma(\mu_x, \nu_z)]$. If we define A'' as $A''(x, z) \iff \exists \sigma[\sigma(\mu_x, \nu_z)]$ then we find ourselves with $\forall x \exists z[A'(x, z)]$, i.e. back to our initial position of wishing to obtain z from x.

In essence, the proof leads to an unending loop of needing to find a z from an x and thus the introduction of the additional sequences prevents the proof. If we did not have additional sequences, there would be no additional zs to find and thus we would be able to proceed as normal.

Let us say for a moment that we somehow surmount this issue and find our ν_z; exactly what information about ν_z would we need? Even if we argue that we can just provide it's law and a list of its elements, we make no provision for it's relationships with μ_x; which ones of those would be required? To have any hope of doing this we would need to understand exactly how we obtain y from σ; but, since we are unable to verify exactly how our $\hat{e}$ works, we have no way of providing a general method to obtain y from σ.

In summary; we can provide no functional or algorithmic way to obtain σ (or $\underline{\nu}$) from x and thus we can only assert classical existence (it is reasonable to exist) rather than constructive existence. Whether the fault lies with the additional forms of information available, or with the additional sequences, or both, is not clear from this example.

Below; we give a slightly stronger version of $BC\text{-}N\text{-}KS$ that omits the $\exists\underline{\nu}$ quantifier.

$BC\text{-}N\text{-}KS^{+}$ $\forall\underline{\mu}\exists\underline{x}[A(\underline{\mu},\underline{x})] \rightarrow \exists\underline{\hat{e}} \in \hat{K}_0\forall\underline{\mu}\exists\sigma[\sigma(\underline{\mu}) \wedge A(\underline{\mu},\underline{\hat{e}}(\sigma))]$

Using a similar method to the conventional theory, we can now obtain the following theorem, a parallel to the derivation of $AC\text{-}NN$ in the conventional theory (see Theorem 2.4.2.5).

Theorem 4.8.1

$BC\text{-}N\text{-}KS^{+} \vdash AC\text{-}NN$

Proof:

Assume $BC\text{-}N\text{-}KS^{+}$ and the hypothesis of $AC\text{-}NN$ ($\forall x\exists y[A(x,y)]$).

Define $\mu \in M_N \iff \exists x[\mu \equiv \lambda z.x]$ (the species of sequences defined by a specific form of constant function).

Given any x define $\mu_x \equiv \lambda z.x$. Clearly, every $\mu \in M_N$ can be constructed in this way.

Define the predicate A' as $A'(\mu, y) \iff A(\mu(0), y)$.

Thus, using these definitions, we can rewrite the hypothesis of $AC\text{-}NN$ as $\forall\mu \in M_N\exists y[A'(\mu,y)]$.

Hence, by $BC\text{-}N\text{-}KS^{+}$ (suitably relativised to the spread M_N), $\exists\hat{e} \in \hat{K}_0$ such that $\forall\mu \in M_N\exists\sigma[\sigma(\mu) \wedge A'(\mu,\hat{e}(\sigma))]$ (i).

Because we know that $\forall\mu \in M_N\exists x[\mu \equiv \mu_x]$, that $A'(\mu,y) \iff A(\mu(0),y)$, the latter part of (i) can be rewritten as $\forall x\exists\sigma[\sigma(\mu_x) \wedge A(x,\hat{e}(\sigma))]$ (ii).

Given any μ and ν in M_N such that $\mu = \nu$, we know by (i) that $\exists\sigma[\sigma(\mu) \wedge \hat{e}(\sigma) \in N]$ and $\exists\sigma[\sigma(\mu) \wedge \hat{e}(\sigma) \in N]$.

If $\sigma \equiv \sigma'$, then $\hat{e}(\sigma) = \hat{e}(\sigma')$.

Assume that $\sigma \not\equiv \sigma'$ and $\hat{e}(\sigma) = y_1 \neq y_2 = \hat{e}(\sigma')$.

Define $\gamma \in M_N$ such that $\mu = \gamma = \nu$ and $(\sigma \sqcap \sigma')(\gamma)$. We can be certain that $\sigma \sqcap \sigma'$ is consistent (that it defines finite information about a choice sequence) because both sequences are extensionally equal and thus it would be impossible for one piece of information to contradict the other as there are no other sequences present.

So $\hat{e}(\sigma \sqcap \sigma') = \Delta$ by $\hat{K}_0 2$; but this would violate $\hat{K}_0 3$ giving us a contradiction!

Thus $\hat{e}$ must be extensional.

Via the fact that $\hat{e}$ is extensional; we can assert that, given an x our σ_x need only contain atoms of the form $SE(i, x)$ and $Law(\lambda z.x)$ as these are the only atoms guaranteed to be present.

Define $\hat{e}'$ in the following way:

Replace every reference in $\hat{e}$ to an atom of the form $SE(i, x)$ with a reference to the atom $SE(0, x)$.

$\hat{e}'$ is clearly extensionally equal to $\hat{e}$ and $\hat{e}'$ is also clearly a member of $\hat{K}_0$.

Given any x we can define $\sigma_x \equiv SE(0, x) \sqcap Law(\lambda z.x)$ and be certain that $\hat{e}'(\sigma_x) \in N$

If we relabel $\hat{e}'$ as $\hat{e}$, we can now rewrite (ii) as $\forall x[\sigma_x(\mu_x) \wedge A(x, \hat{e}(\sigma_x))]$ (iii).

Define the function f as $f \equiv \lambda x.\hat{e}'(\sigma_x)$.

Thus by rewriting (iii) we have $\exists f \forall x[A(x, f(x))]$, and hence, by discharging our assumptions, $\forall x \exists y A(x, y) \rightarrow \exists f \forall x A(x, f(x))$ as required. ♠

This shows that extending the bounds of our finite information isn't what prevents the proof of $AC\text{-}NN$, it is the inclusion of the additional sequences $\underline{\nu}$. This highlights just how strong an assumption we make in the conventional interpretations of $BC\text{-}N$; however, some ground can be made using the theory of knowledge states as we now allow additional intensional information about our sequences to be present (and indeed, used).

§4.9. Resolving Fletcher's Paradoxes

In the earlier literature review, a major flaw with the conventional theory was the occurrence of Fletcher's paradoxes. In this section, we shall show the point at which the construction of each of Fletcher's paradoxes fail in the new theory.

Fletcher's first paradox is resolved as follows in the extended theory.

> Define the choice sequences $\mu \in s_{\{0,1\}}$ and $\nu \in s_{\{0,1\}}$ such that $\forall x[\nu(x) = 1 - \mu(x)]$, and hence $\forall x[\mu(x) = 1 - \nu(x)]$.
>
> Also, our generalised open data only states that 'given any predicate on tuples of choice sequences A we have $\forall \underline{\mu}(A(\underline{\mu}) \to \exists \sigma(\sigma(\underline{\mu}) \wedge \forall \underline{\nu}(\sigma(\underline{\nu}) \to A(\underline{\nu}))))$ where A has no other free choice sequence variables save $\underline{\mu}$'. It in no way enforces us to just rely on the elements of a sequence, nor does it state that σ is unique or forbid relations between sequences.
>
> In our theory the notion of a 'lawless sequence' is slightly different ($KS\text{-}LL$ rather than $KT\text{-}LL$). For μ and ν this is defined as $\forall \sigma[\sigma(\mu) \to \sigma \in \Sigma_{SE}]$ and $\forall \sigma[\sigma(\nu) \to \sigma \in \Sigma_{SE}]$. The reader will immediately notice that $\forall \sigma[\sigma(\mu, \nu) \to \sigma \in \Sigma_{SE}]$ is **false** as it is clear that $(LR(R, 2, 1) \sqcap LR(R, 1, 2))(\mu, \nu)$, where R is the relation $\forall x[\{w_2\}(x) = 1 - \{w_1\}(x)]$, holds. Thus we are in a position to assert that both μ and ν may be $KS\text{-}LL$ without contradiction, even if the relation $\forall x[\mu(x) = 1 - \nu(x)]$ holds.
>
> In essence, the paradox is avoided by two key facts: firstly, because our generalised open data allows additional forms of finite information (specifically relations between

sequences), and secondly, because our notion of lawlessness allows for relations between choice sequences.

This leads us to Fletcher's second paradox, and its resolution, below.

Given a lawless choice sequence $\mu \in s_{\{0,1\}}$, define $\nu \in s_{\{0,1\}}$ such that $\forall x > 1[\nu(x+1) = \mu(x)] \wedge \nu(1) = 0$.

If both sequences were lawless (KT-LL) then they would satisfy the condition given above. However, this is clearly not the case as can be seen by applying conventional open data with $A(\mu, \nu)$ as $\forall x[\nu(x+1) = \mu(x)]$. Clearly, this is true but it is not decided via initial segments alone and hence we have a contradiction. In the extended theory, we do not arrive at this contradiction as open data is not just valid for lawless (KS-LL) sequences and our definition of lawlessness allows there to be such relationships between sequences.

Fletcher's third paradox is resolved below.

This paradox relies on the principle $\mu = \nu \iff \mu \equiv \nu$, which we have rejected as we can prove its contradictoriness with a simple counter-example of μ and ν where $(Law(1, \lambda x.x) \sqcap Law(2, \lambda x.\frac{2x}{2})(\mu, \nu)$. Without this principle the paradox is impossible to construct. Even if we were to try and construct it in another way, we would first need a notion of 'proto-lawlessness', which is impossible in our theory as the distinction between lawless and proto-lawless in the conventional theory relies on the distinction between restriction and generation, which we do not make!

Finally, we reach Fletcher's fourth paradox and provide our resolution of it below.

A key notion in the conventional theory is that of a proto-lawless sequence and this occurs because the conventional theory differentiates between generating an element and restricting it before generation. We lack a notion of proto-lawlessness because we do not make this differentiation. Thus, without a notion of proto-lawlessness (the closest we have is the 'proto-lawless' (empty) knowledge state $\sigma_\emptyset$), we have no fourth paradox in its original form. The closest we can come is the second paradox, which has already been resolved by our generalised open data!

This concludes our analysis of Fletcher's paradoxes in the extended theory. They no longer occur and I feel that this is reasonable grounds to assert that our notion of lawlessness improves upon both KT-LL and M-LL. KS-LL's evasion of Fletcher's paradoxes and its wholly decidable nature (under the one manual rule) gives it a distinct edge over both KT-LL and M-LL.

§4.10. General Bar Induction

The final notion we require for analysis is a generalised notion of bar induction. Before we can construct our bar induction schema, we must first discuss about the intensionality or extensionality of bars and what this means for any predicate we use.

In our extended theory, a *bar* R is a property that eventually holds for some finite information of every choice sequence. More formally, we write that R *bars the spread* s *iff* $\forall \underline{\mu} \in s \exists \sigma \in \Sigma_{SE}[\sigma(\underline{\mu}) \wedge R(\sigma)]$, and we write that R *bars the universal spread* iff $\forall \underline{\mu} \exists \sigma \in \Sigma_{SE}[\sigma(\underline{\mu}) \wedge R(\sigma)]$.

Given any choice sequence, the only meaningful information we can guarantee to be present are finite conjunctions of SE atoms. The bar $R(\sigma) \iff \exists s[Spread(s) \in \sigma]$ definitely demonstrate that some intensional information could be put to use in defining a meaningful bar; however, it is difficult to see how one could work with such information in any useful way given that all such Rs would be vacuously true for a certain atom. The complexity of ensuring we have defined an upward hereditary property for each of these forms of knowledge leads us to reduce the scope of our bar induction to only barring Baire space (which the conventional theory shows is sufficient for analysis), and it is from here that we make our first assumption below about our bars (R).

BI-$A1$ $\forall \underline{\mu} \in s \exists \sigma \in \Sigma_{SE}[\sigma(\mu) \wedge R(\sigma)]$

Like the conventional notions of bar induction exposited in §2.5.1, the assertion above is actually slightly stronger than just asserting that R is extensional, because we are actually asserting that R is 'graph extensional' (see §2.4.1).

A point the author feels compelled to make is that the intensional R in the paragraph above BI-$A1$ does indicate that it may be possible to construct a stronger schema of bar induction using the language of knowledge states; for now we will limit ourselves to obtaining a notion sufficient to follow the conventional path to analysis. The derivation of a stronger schema of bar induction, or at least a full exploration of the feasibility of such a schema, would be a worthy project for the future.

In bar induction we usually assert $\forall \sigma[R(\sigma) \to A(\sigma)]$; combined with BI-$A1$ this means that the only σs R 'passes' to A will be in Σ_{SE}. This means that our As only have to be upward hereditary over Baire Space (since they will only ever be dealing with SE atoms), though in a way different from the conventional notions of bar induction explored in §2.5.1. The example that best illustrates this is given below.

Example 4.10.1

Define $R(\sigma) \iff \exists y[SE(0, 4, y) \in \sigma]$, i.e. a choice sequence hits the bar when we know its 5^{th} element; clearly, every choice sequence will eventually have a 5^{th} element (and we can see this by having that element presented to us as a Σ_{SE} knowledge state hence satisfying BI-$A1$) and we may also be certain if we have the 5^{th} element (or not), i.e. the bar is decidable.

Define A as the predicate 'the choice sequence with known elements σ has a 5^{th} element'.

Clearly, $R(\sigma) \rightarrow A(\sigma)$ by definition; however, this leads to the case where we have $\forall y[A(SE(0,4,y)]$ ('any sequence where we know the 5^{th} element has a 5^{th} element').

We would like to be able to make an inductive step up the tree; however, we only have $\forall y[A(\sigma_\emptyset \sqcap SE(0,4,y))]$. We would like to imply $A(\sigma_\emptyset)$ from this, and hence the logical property to insist upon for A is $\exists w \exists x \forall y[A(\sigma \sqcap SE(w,x,y))] \rightarrow A(\sigma)$.

This is where our assumption on the extensionality of R pays dividends; if we had not insisted on purely extensional bars, then we would also require additional powers/restrictions on A based on the various limitations on R. As things stand, we find ourselves in a desirable position of knowing that this is the only type of upward hereditary we need to consider for now.

Hence, in general, we assert that A must satisfy the following.

*BI-A*2 $\forall\sigma[\forall\sigma \in \Sigma_{SE}[\exists w \exists x \forall y[(\sigma \sqcap SE(w,x,y)) \in \Sigma_{SE} \wedge A(\sigma \sqcap SE(w,x,y))] \rightarrow A(\sigma)]]$

We invite the reader to consider that this clause is no weaker than the one in the conventional theory, which also only looks at all possible values of one element of a sequence. It is no stronger either, though perhaps it could be argued that it is 'smarter'; if R requires only two elements to be satisfied, it does not matter if these are the third or thousandth, our notion would only necessitate a tree (see §1.2) two deep, and hence, two steps back. On the other hand, the conventional theory would require a very variable number of steps back, from three to a thousand, a much less efficient method.

Thus, with these ideas in mind, our two schemas of bar induction are given below.

BI_D-KS

$$\{\forall \underline{\mu} \exists \sigma \in \Sigma_{SE}[\sigma(\underline{\mu}) \wedge R(\sigma)]$$

$$\wedge \forall \sigma \in \Sigma_{SE}[R(\sigma) \vee \neg R(\sigma)]$$

$$\wedge \forall \sigma \in \Sigma_{SE}[R(\sigma) \rightarrow A(\sigma)]$$

$$\wedge \forall \sigma \in \Sigma_{SE}[\exists x \exists w \forall y[(\sigma \sqcap SE(w,x,y)) \in \Sigma_{SE} \wedge A(\sigma \sqcap SE(w,x,y))] \rightarrow A(\sigma)]\}$$

$$\rightarrow A(\sigma_{\emptyset})$$

BI_M-KS

$$\{\forall \underline{\mu} \exists \sigma \in \Sigma_{SE}[\sigma(\underline{\mu}) \wedge R(\sigma)]$$

$$\wedge \forall \sigma \in \Sigma_{SE} \forall w \forall x \forall y[R(\sigma) \rightarrow (R(\sigma \sqcap SE(w,x,y)]$$

$$\wedge \forall \sigma \in \Sigma_{SE}[R(\sigma) \rightarrow A(\sigma)]$$

$$\wedge \forall \sigma \in \Sigma_{SE}[\exists x \exists w \forall y[(\sigma \sqcap SE(w,x,y)) \in \Sigma_{SE} \wedge A(\sigma \sqcap SE(w,x,y))] \rightarrow A(\sigma)]\}$$

$$\rightarrow A(\sigma_{\emptyset})$$

The reader may wonder why the fourth condition both both schemata is not $\forall \sigma \in \Sigma_{SE}[\exists w \forall x \forall y[(\sigma \sqcap SE(w,x,y)) \in \Sigma_{SE} \rightarrow A(\sigma \sqcap SE(w,x,y))]$. The example below offers clarity as to why we do this:

Example 4.10.2

Define $A(\sigma) \iff \forall \mu[\sigma(\mu) \rightarrow \mu(99)$ *is odd or even*$]$; i.e A holds for σ iff any sequence consistent with knowledge state σ has a 100^{th} element.

So all we need to assert $A(\sigma_{\emptyset})$ is to know that $\forall x[A(SE(99,x))]$ (that no matter what value the 100^{th} element is it satisfies A) as the order in which our information is given to us shouldn't matter for a graph extensional predicate. Taking the time to also assert $\forall x[A(SE(0,x))]$, $\forall x[A(SE(1,x))]$, ... would be a pointless redundancy of the kind we wish to avoid.

In essence, because the order in which we are given elements should not matter (remember, our predicates are graph extensional), we need only consider one possible order in which the elements are given to us. Checking additional orders would be pointless. However, should a fully generalised notion of bar induction ever be constructed where R and A need not be graph extensional, then our fourth clause may well need to consider these alternative orderings (see §7.3.4 for more details on this).

We will now proceed to provide a very familiar result below.

Theorem 4.10.1

$BI_M\text{-}KS \vdash BI_D\text{-}KS$ (Generalisation of theorem 2.5.1.1)

Proof:

Assume $BI_M\text{-}KS$ holds and the premises of $BI_D\text{-}KS$ hold for a bar R and a predicate A.

Define R' as $\forall\sigma \in \Sigma_{SE}[R'(\sigma) \iff \exists\sigma' \in \Sigma_{SE}[\sigma' \subseteq \sigma \to R(\sigma')]]$.

Define A' as $\forall\sigma \in \Sigma_{SE}[A'(\sigma) \iff \exists\sigma' \in \Sigma_{SE}[\sigma' \subseteq \sigma \to A(\sigma')] \vee R'(\sigma)]$.

R' is monotonic by definition.

By the assumption $\forall\underline{\mu}\exists\sigma \in \Sigma_{SE}[\sigma(\underline{\mu}) \wedge R(\sigma)]$ we have $\forall\underline{\mu}\exists\sigma \in \Sigma_{SE}[\sigma(\underline{\mu}) \wedge R'(\sigma)]$.

By the definition of A' we have $\forall\sigma \in \Sigma_{SE}[R'(\sigma) \to A'(\sigma)]$.

Given any arbitrary σ suppose we have x_0 and w_0 such that $\forall y[(\sigma \sqcap SE(w_0, x_0, y)) \in \Sigma_{SE} \wedge A'(\sigma \sqcap SE(w_0, x_0, y))]$.

R' is decidable by its construction, so we can argue our final point by cases.

If $R'(\sigma)$ then $A'(\sigma)$ by definition.

If $\neg R'(\sigma)$, then we have $\forall y[R'(\sigma \sqcap SE(w_0, x_0, y)) \rightarrow (R(\sigma \sqcap SE(w_0, x_0, y)) \vee \exists \sigma' \subset \sigma[R(\sigma' \sqcap SE(w_0, x_0, y)) \wedge \neg R(\sigma')])]$.

If the first disjunct, then we know that $\forall y[R(\sigma \sqcap SE(w_0, x_0, y)) \rightarrow A(\sigma \sqcap SE(w_0, x_0, y))]$, so $\forall y[R'(\sigma \sqcap SE(w_0, x_0, y)) \rightarrow A(\sigma \sqcap SE(w_0, x_0, y))]$. Since A is upward hereditary we obtain $A(\sigma)$, and hence $A'(\sigma)$ as required.

If the second disjunct, then we know that $\forall y[R(\sigma' \sqcap SE(w_0, x_0, y))]$, and hence we get $\forall y[A(\sigma' \sqcap SE(w_0, x_0, y))]$. As A is upward hereditary we obtain $A(\sigma')$ and hence, since $\sigma' \subseteq \sigma$, we obtain $A'(\sigma)$ as required.

Thus, we have $\forall \sigma[\exists x \exists w \forall y[A'(\sigma \sqcap SE(w, x, y))] \rightarrow A'(\sigma)]$ in both cases.

By BI_M we can imply $A'(\sigma_\emptyset)$, and hence $A(\sigma_\emptyset)$ as required. ♠

It is now that we begin to see the price of our weaker BC-N-KS; it no longer offers us a generalisation of theorem 2.5.1.2. To illustrate how this comes about, we offer below the following proof of BI_D-$KS \wedge BC$-N-$KS \vdash BI_M$-KS up until it breaks down, and we explain why it now fails.

Assume that BI_D-KS holds and that the bar R and the predicate A satisfy the hypothesis of BI_M-KS.

By the assumptions of BI_M-KS we have that,

$\forall \underline{\mu} \exists \sigma \in \Sigma_{SE}[\sigma(\underline{\mu}) \wedge R(\sigma)]$ (i)

$\forall \sigma \in \Sigma_{SE} \forall w \forall y \forall x[R(\sigma) \rightarrow R(\sigma \sqcap SE(w, x, y))]$ (ii)

(i) and (ii) allow us to say that given our σ, we can construct from this a $\sigma' \in \Sigma_{IS}$, such that $\sigma'(\mu) \wedge R'(\sigma)$ by 'filling in the gaps' by querying μ for the missing elements.

Define R^{is} as follows $R^{is}(n) \iff \exists\sigma \in \Sigma_{IS}(\sigma \sim n \wedge R(\sigma))$ (iii) (for the definition of $\sigma \sim n$ see §4.6).

By (i) and (iii) we can rewrite $\forall\mu\exists\sigma \in \Sigma_{SE}[\sigma(\mu) \wedge R(\sigma)]$ as $\forall\mu\exists x[R^{is}(\bar{\mu}(x))]$.

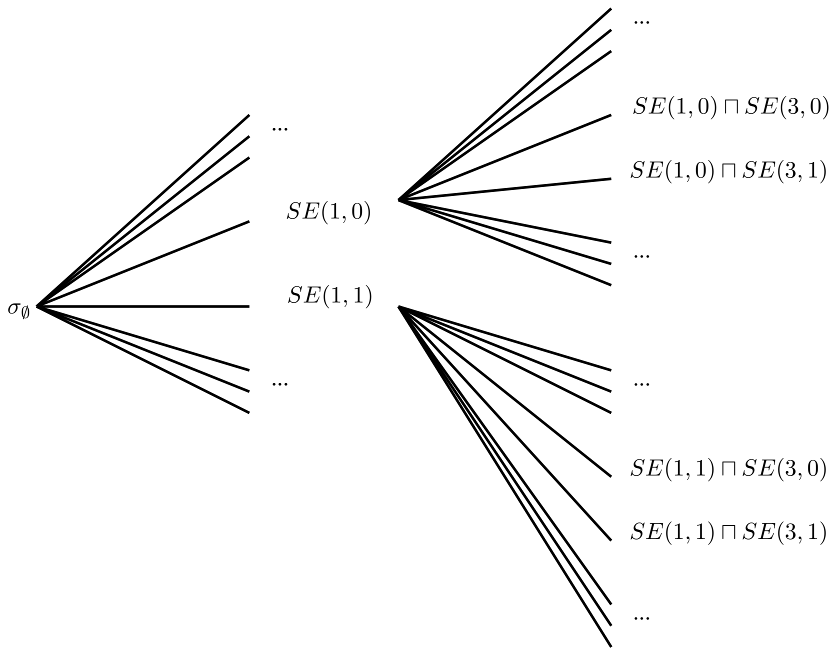
Hence, by *BC-N-KS* we have that $\exists\hat{e} \in \hat{K}_0\forall\mu\exists\sigma[\exists\underline{\nu}[\sigma(\underline{\mu},\underline{\nu})] \wedge R^{is}(\bar{\mu}(\underline{\hat{e}}(\sigma)))]$.

If we were to have $\exists\hat{e} \in \hat{K}_0\forall\mu\exists\sigma[\sigma(\underline{\mu}) \wedge R^{is}(\bar{\mu}(\underline{\hat{e}}(\sigma)))]$ instead of the line above, then we could have argued that the extensionality of R forces σ to be in Σ_{SE}; however, we can make no such argument for the case containing $\underline{\nu}$ and it is at this point that the proof breaks down. This does not rule out that another method of proof exists, merely that this one fails and strongly indicates that *BC-N-KS* is not quite strong enough to obtain this result. This, however, is no slur on *BC-N-KS* as our notion of bar induction is not truly generalised whereas *BC-N-K* is; this we feel is the root of the failure of the proof and not any inadequacy in our notion of BI_D-KS.

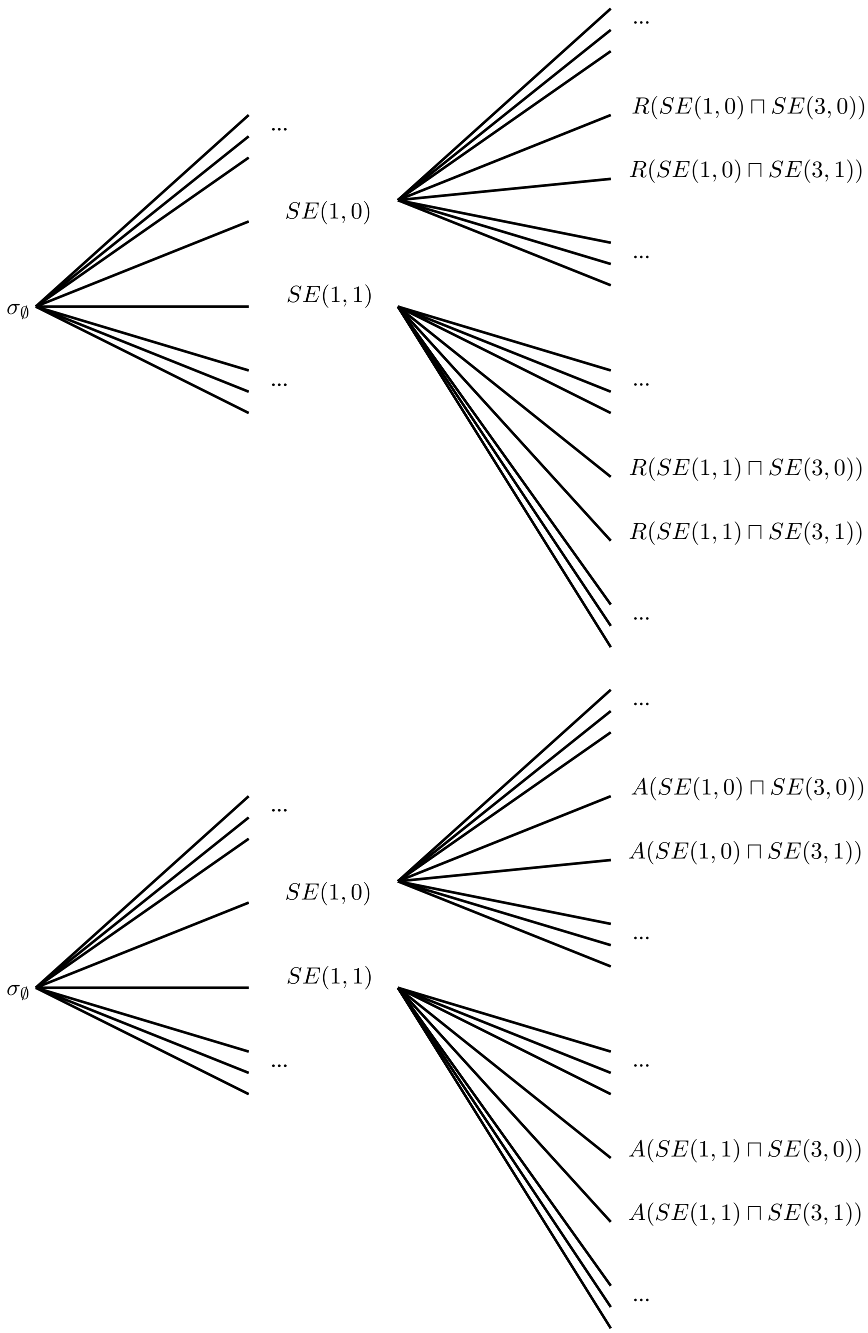
These notions of bar induction make clear sense; however, relating them to the conventional notions of bar induction is problematic. While it is true that $BI_D \rightarrow BI_D$-KS, the converse does not appear to hold true; the same holds true for $BI_M \rightarrow BI_M$-KS. Below we highlight something of interest.

Given the way our bars are formulated, it is clear that they simply need a **finite** list of elements. Let's define a bar; $R(\sigma) \iff \exists x[SE(1,x) \in \sigma] \wedge \exists x[SE(3,x) \in \sigma]$, i.e. we hit the bar when we have the 2^{nd} and 4^{th} elements of our sequence. For simplicity's sake, we'll also restrict ourselves to a binary fan; the argument easily generalises.

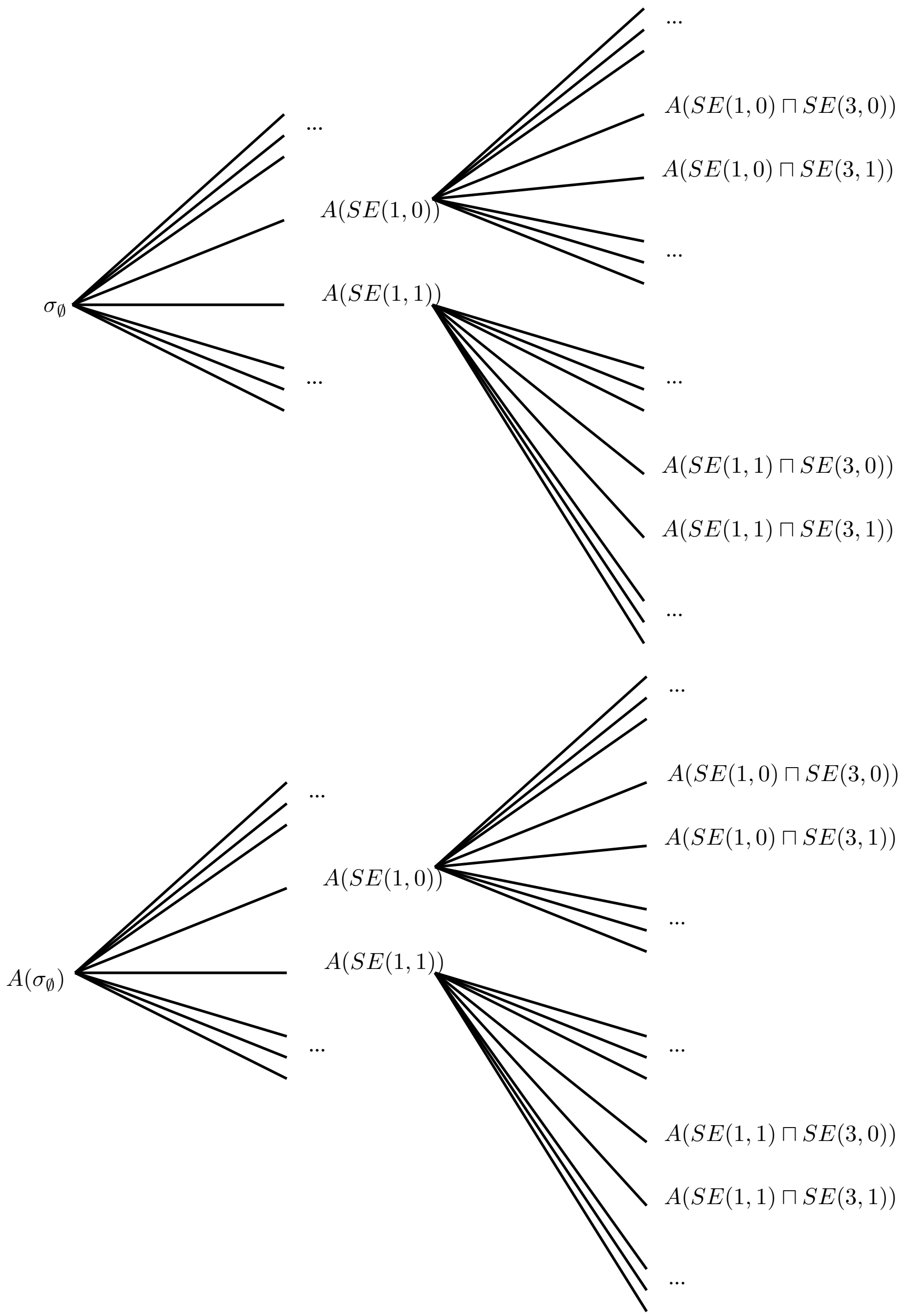
Figure 3 represents the two 'shortest' paths to hit the bar; surprisingly these both generate enough As via $R(\sigma) \rightarrow A(\sigma)$ to imply $A(\sigma_\emptyset)$ when considered alone as seen in figures 4 and 5.



$Fig\ 3\ :\ Shortest\ paths$



$Fig\ 4\ :\ Path\ to\ A(\sigma_\emptyset)\ 1$



$Fig\ 5\ :\ Path\ to\ A(\sigma_\emptyset)\ 2$

This need of only one sufficient path (one of which will always be in Σ_{IS} through the simple rearrangement of the atoms) leads to the interesting thought below.

> De jure we require, for any $\sigma \in \Sigma_{SE}$, $\exists x \forall y[A(\sigma \sqcap SE(x,y))]$ to step back, but de facto we can always make use of , for any $\sigma \in \Sigma_{SE}$, $\exists x \geq MEI(\sigma) \forall y[A(\sigma \sqcap SE(x,y))]$, where we remind the reader that $MEI(\sigma)$ denotes the largest element index present in σ.

That this trick will work for any of our Rs is given, and thus we propose the following alternative schema of decidable bar induction which, while less general, is sufficiently strong enough to derive BI_D.

BI_D-KS-ALT

$$\begin{aligned}
&\{\forall \underline{\mu} \exists \sigma \in \Sigma_{SE}[\sigma(\underline{\mu}) \wedge R(\sigma)] \\
&\wedge \forall \sigma \in \Sigma_{SE}[R(\sigma) \vee \neg R(\sigma)] \\
&\wedge \forall \sigma \in \Sigma_{SE}[R(\sigma) \rightarrow A(\sigma)] \\
&\wedge \forall \sigma \in \Sigma_{SE}[\exists x \geq MEI(\sigma) \exists w \forall y[A(\sigma \sqcap SE(w,x,y))] \rightarrow A(\sigma)]\} \\
&\quad \rightarrow A(\sigma_\emptyset)
\end{aligned}$$

This version is slightly stronger in a formal sense; however, it is no stronger in an informal sense, as argued above. From this we obtain the following theorem.

Theorem 4.10.2

BI_D-KS-$ALT \vdash BI_D$

Proof:

If given some $\sigma \in \Sigma_{IS}$, we define $|\sigma|$ to be $|n|$ for the n such that $n \sim \sigma$ then BI_D is clearly equivalent to the following.

$$\{\forall \mu \exists \sigma \in \Sigma_{IS}[\sigma(\mu) \wedge R(\sigma)]$$

$$\wedge \forall \sigma \in \Sigma_{IS}[R(\sigma) \vee \neg R(\sigma)]$$

$$\wedge \forall \sigma \notin \Sigma_{IS}[\neg R(\sigma)](*)$$

$$\wedge \forall \sigma \in \Sigma_{IS}[R(\sigma) \rightarrow A(\sigma)]$$

$$\wedge \forall \sigma \in \Sigma_{IS}[\forall y[A(\sigma \sqcap SE(0, |\sigma|, y))] \rightarrow A(\sigma)]\}$$

$$\rightarrow A(\sigma_{\emptyset})$$

Assume R and A satisfy the above hypotheses.

Define R' as $\forall \sigma \in \Sigma_{SE}[R'(\sigma) \iff R(\sigma)]$.

Define A' as $\forall \sigma \in \Sigma_{SE}[A'(\sigma) \iff A(\sigma) \vee R'(\sigma)]$.

R' is decidable by definition.

By the assumption $\forall \mu \exists \sigma \in \Sigma_{IS}[\sigma(\mu) \wedge R(\sigma)]$ we have $\forall \mu \exists \sigma \in \Sigma_{SE}[\sigma(\underline{\mu}) \wedge R'(\sigma)]$ since $\Sigma_{IS} \subset \sigma_{SE}$.

By the definition of A' we have $\forall \sigma \in \Sigma_{SE}[R'(\sigma) \rightarrow A'(\sigma)]$.

Given any arbitrary σ suppose we have that, we have, for some $x_0 \geq MEI(\sigma)$, that $\forall y[A'(\sigma \sqcap SE(x, y))]$.

R' is decidable by its construction so we can argue our final point by cases.

If $R'(\sigma)$ then $A'(\sigma)$ by definition.

If $\neg R'(\sigma)$ then we have $\forall y[R'(\sigma \sqcap SE(x_0, y)) \iff R(\sigma \sqcap SE(x_0, y))]$.

By (*) and $(\sigma \sqcap SE(x_0, y)) \in \Sigma_{IS}$ we know that $x_0 \leq MEI(\sigma)$, but $x_0 \geq MEI(\sigma)$ by hypothesis so $\forall y[A(\sigma \sqcap SE(MEI(\sigma), y))]$ which, by hypothesis implies $A(\sigma)$ and hence $A'(\sigma)$ by definition.

Hence, discharging our hypothesis $\forall\sigma \in \Sigma SE[\exists x\forall y[A'(\sigma \sqcap SE(x,y))] \to A'(\sigma)]$ and thus R' and A' satisfy the hypothesis for BI_D-KS-ALT and hence $A'(\sigma_\emptyset)$.

Hence we have either $R'(\sigma_\emptyset) \iff R(\sigma_\emptyset) \to A(\sigma_\emptyset)$ or $A(\sigma_\emptyset)$. Either way we have $A(\sigma_\emptyset)$.

Hence, discharging our assumptions on R and A, we obtain BI_D as required. ♠

While our notion of BI_D-KS-ALT is weaker than BI_M-KS it is still sufficiently strong enough to give us the results we desire for analysis, as all that is required is BI_D. We have yet to obtain a notion of BI_M-KS-ALT that proves BI_M.

5. A New Formal System

§5.1. The System $FIM\text{-}KS$

Before we proceed to construct our formal system, we must first settle upon a formal notion of "lawlike function". This notion is one that has eluded attempts to formally define it in the existing literature so far, and the author is unable to provide any solid refinement in this area to date. Instead, we shall restrict ourselves to the (smallest possible) subspecies of the lawlike functions we may always recognise in a decidable way; i.e. the primitive recursive functions. The reader may prefer a different **decidable** species of total functions for their lawlike functions, and we invite them to use these in place of primitive recursive functions should they so desire. This does mean that our system doesn't allow us to recognise every possible lawlike sequence as lawlike; however, all such unrecognised lawlike sequences (i.e. any constructive sequence that is not primitive recursive) is still quantified over in our choice sequence variables. We are just unable to use the fact that these unrecognised lawlike sequences are lawlike, and must treat them as we do any arbitrary choice sequence. This has no impact on our reduction to analysis in §6. We are now ready to formally construct our system for analysis, $FIM\text{-}KS$, as an extension of $PrAn$.

We begin by extending the language of $PrAn$ to include variables for knowledge states ('σ'), knowledge state connectives ('$\sqcap$', '$\sqcup$'), intensional equality ('$\equiv$'), the ordering on knowledge states ('$\subseteq$') and a second type of function of varying arity ($\hat{e}^i$). '$\sigma_\emptyset$', 'MEI', and Spr are additional constants.

$FIM\text{-}KS\text{-}T7a$ $\sigma_\emptyset$ is a **knowledge term**.

$FIM\text{-}KS\text{-}T7b$ Any knowledge state variable σ is a **knowledge term**.

$FIM\text{-}KS\text{-}T8$ Any function variable $\hat{e}^i$ is a $\hat{K}_0$ **term**.

$FIM\text{-}KS\text{-}T9a$ Given numerical terms t_1, t_2 and t_3, $SE(t_1, t_2, t_3)$ is a **knowledge term**.

$FIM\text{-}KS\text{-}T9b$ Given any primitive recursive functional term T and any numerical term t, $law(t, T)$ is a **knowledge term**.

$FIM\text{-}KS\text{-}T9c$ Given any primitive recursive functional term T such that $Spr(T) = 0$, and any numerical term t, $Spread(t, s)$ is a **knowledge term**.

$FIM\text{-}KS\text{-}T9d$ Given any primitive recursive functional term T_1 such that $Spr(T_1) = 0$, any primitive recursive functional term T_2, and any numerical term t, $Fan(t, T_1, T_2)$ is a **knowledge term**.

$FIM\text{-}KS\text{-}T9e$ Given any $\hat{K}_0$ term T and any tuple of numerical terms t_1, t_2,..., t_x, $LR(\lambda x.T, t_1, t_2, ..., t_x)$ and $SR(\lambda x.T, t_1, t_2, ..., t_x)$ are **knowledge terms**.

$FIM\text{-}KS\text{-}T10$ Given two knowledge terms t and t' then $\ulcorner t \sqcap t' \urcorner$ and $\ulcorner t \sqcup t' \urcorner$ are **knowledge terms**.

$FIM\text{-}KS\text{-}T11$ Given a knowledge term t and a $\hat{K}_0$ term T, then $\ulcorner T(t) \urcorner$ is an **extended numerical term**.

$FIM\text{-}KS\text{-}F$ Any $\hat{K}_0$ function defined by Church's λ (of the form $\lambda x.\hat{e}$ where $\hat{e}$ is a $\hat{K}_0$ function variable) is a $\hat{K}_0$ **term**.

We pause to mention to the reader that the language regarding $\hat{K}_0$ functions is extensible as will be seen in §6.4.

The function MEI is of type $\Sigma \to N^*$ and is defined as below.

$$MEI(\sigma) = \begin{cases} max(x \mid SE(w, x, y) \in \sigma) \ if \ \sigma \in \Sigma_{SE} \\ \nabla \ otherwise \end{cases}$$

The function Spr is of type $M \to N$ and is defined as below.

$$Spr(\mu) = \begin{cases} 0 \ if \ \forall n[\mu(n) = 0 \vee \mu(n) = 1] \wedge \mu(\langle\rangle) = 0 \wedge \forall n[\mu(n) = 0 \iff \exists x[\mu(n * x) = 0]] \\ 1 \ otherwise \end{cases}$$

We extend the formula formation of $PrAn$ such that the atomic formula of the language are of the one of forms given below.

$\ulcorner t = t' \urcorner$ where t and t' are both numerical terms, function terms or $\hat{K}_0$ terms.

or

$\ulcorner t \equiv t' \urcorner$, where t and t' are both knowledge terms, functional terms or $\hat{K}_0$ terms.

or

$\ulcorner t(t') \urcorner$, where t is a knowledge term and t' is a functional term.

Further formulae are constructed from these atoms utilising the logical connectives and quantifiers as per usual.

Because we have two types of equality, we introduce the following axiom to link them.

FIM-KS-EQ $\ulcorner t \equiv t' \urcorner \rightarrow \ulcorner t = t' \urcorner$ where t and t' are functional terms or $\hat{K}_0$ terms.

We define some useful meta-notational concepts below.

We will denote any primitive recursive functional term of type $N^x \mapsto N$, which contains no choice sequence parameters, with the special symbols f and g.

Given a knowledge term t, we will write $\ulcorner t \in \Sigma_{SE} \urcorner$ to mean that t satisfies the conditions to belong to Σ_{SE} as outlined in (§4.6).

Likewise, when we write $\ulcorner t \in \Sigma_{IS} \urcorner$ we mean that t satisfies the conditions to belong to Σ_{IS} as outlined in (§4.6).

We include the axioms for knowledge states (justified in §4.2) below.

AX-MOD $\forall\sigma\forall\underline{\mu}[\sigma(\underline{\mu}) \rightarrow |\sigma| \leq |\underline{\mu}|\ ]$

σ-μ-1 $\forall\underline{\mu}[A(\underline{\mu}) \rightarrow \exists\sigma[\exists\underline{\nu}[\sigma(\underline{\mu},\underline{\nu})] \wedge A'(\sigma)]]$, where A has no other free choice sequence variables and $|\underline{\mu}|$ need not equal $|\underline{\nu}|$.

σ-μ-2 $\forall\sigma[A'(\sigma) \rightarrow \forall\underline{\mu}[\exists\underline{\nu}[\sigma(\underline{\mu},\underline{\nu})] \rightarrow A(\underline{\mu})]]$, where A has no other free choice sequence variables and $|\underline{\mu}|$ need not equal $|\underline{\nu}|$.

σ-*sem*-1 $\forall w \forall x \forall y \forall \mu_0, ...\mu_i[(SE(w,x,y))(\mu_0,...\mu_i) \iff \mu_w(x) = y]$

σ-*sem*-2 $\forall w \forall s_{Spread(s)} \forall \mu_0, ...\mu_i[(Spr(w,s))(\mu_0,...\mu_i) \rightarrow \mu_w \in s]$

σ-*sem*-3 $\forall w \forall s \forall f \forall \mu_0, ...\mu_i[(Fan(w,s,f))(\mu_0,...\mu_i) \rightarrow \mu_w \in s \wedge \forall x[\mu_w(x) \leq f(\bar{\mu_w}(x))]]$

σ-*sem*-4 $\forall R \forall w \forall y_0, ..., y_k \forall \mu_0, ...\mu_i[(LR(R,w,y_0,...,y_i))(\mu_0,...\mu_i) \rightarrow \forall j \leq k[w \neq y_j] \wedge \forall x[\mu_w(x) = R(\{y_0\},...,\{y_k\})]]$

σ-*sem*-5 $\forall R \forall y_0, ..., y_k \forall \mu_0, ...\mu_i[(SR(R,y_0,...,y_k))(\mu_0,...\mu_i) \rightarrow R(\{y_0\},...,\{y_k\})]$

(We quickly remind the reader that we cannot always construct a **decidable** A' from A and we say A' is not always constructible to mean this).

The axioms for orderings of knowledge state are introduced below (justified in §4.5).

O-1a $\forall \sigma_\epsilon[\sigma_\epsilon \subseteq \sigma_\epsilon]$

O-1b $\sigma_\emptyset \subseteq \sigma_\emptyset$

O-2a $\sigma_1 \subseteq \sigma_2 \rightarrow \sigma_1 \subseteq (\sigma_2 \sqcap \sigma_3)$

O-2b $\sigma_1 \subseteq \sigma_3 \rightarrow \sigma_1 \subseteq (\sigma_2 \sqcap \sigma_3)$

O-3 $(\sigma_1 \subseteq \sigma_3) \wedge (\sigma_2 \subseteq \sigma_3) \rightarrow (\sigma_1 \sqcap \sigma_2) \subseteq \sigma_3$

O-4a $\sigma_1 \subseteq \sigma_3 \rightarrow (\sigma_1 \sqcup \sigma_2) \subseteq \sigma_3$

O-4b $\sigma_2 \subseteq \sigma_3 \rightarrow (\sigma_1 \sqcup \sigma_2) \subseteq \sigma_3$

O-5 $(\sigma_1 \subseteq \sigma_2) \wedge (\sigma_1 \subseteq \sigma_3) \rightarrow \sigma_1 \subseteq (\sigma_2 \sqcup \sigma_3)$

O-6 $\subseteq$ is the minimum relation satisfying O-1 to O-5

C-1 $\sigma_1 \cong \sigma_2 \iff \sigma_1 \subseteq \sigma_2 \wedge \sigma_2 \subseteq \sigma_1$

We enforce the following axioms on our $\hat{K}_0$ functions.

$\hat{K}_0 1$ $\forall \mu_0, \mu_1, ..., \mu_j[j \geq i \rightarrow \exists \sigma[\exists \nu_0, ..., \nu_{j'}[\sigma(\mu_0, \mu_1, ..., \mu_j, \nu_0, ..., \nu_{j'})] \wedge \hat{e}^i(\sigma) \in N]]$ (totality)

$\hat{K}_0 2$ $\forall\sigma\forall\sigma'[\sigma \subseteq \sigma' \rightarrow \hat{e}^i(\sigma) \preceq \hat{e}^i(\sigma')]$ (monotonicity)

$\hat{K}_0 3$ $\forall\sigma[\exists\underline{\mu}[\sigma(\underline{\mu})] \rightarrow \hat{e}^i(\sigma) \prec \Delta]$ (contradictoriness)

Our schema of bar induction is BI_D-KS-ALT (justified in §4.10) given below.

BI_D-KS-ALT

$$\{\forall\underline{\mu}\exists\sigma \in \Sigma_{SE}[\sigma(\underline{\mu}) \wedge R(\sigma)]$$

$$\wedge \forall\sigma \in \Sigma_{SE}[R(\sigma) \vee \neg R(\sigma)]$$

$$\wedge \forall\sigma \in \Sigma_{SE}[R(\sigma) \rightarrow A(\sigma)]$$

$$\wedge \forall\sigma \in \Sigma_{SE}[\exists x \geq MEI(\sigma)\exists w\forall y[A(\sigma \sqcap SE(w,x,y))] \rightarrow A(\sigma)]\}$$

$$\rightarrow A(\sigma_\emptyset)$$

Our continuity axiom is BC-N-KS (justified in §4.8) as given below.

BC-N-KS $\forall\underline{\mu}\exists\underline{x}[A(\underline{\mu},\underline{x})] \rightarrow \exists\underline{\hat{e}}\forall\underline{\mu}\exists\sigma[\exists\underline{\nu}[\sigma(\underline{\mu},\underline{\nu})] \wedge A(\underline{\mu},\underline{\hat{e}}(\sigma))]$ where A has no other free choice sequence variables.

Due to the fact that we have been unable to provide a proof of AC-NN from BC-N-KS we also find ourselves forced to include an axiom of choice, given below.

AC-NN $\forall x\exists y[A(x,y)] \rightarrow \exists f\forall x[A(x,f(x))]$

This concludes the construction of FIM-KS. To summarise, FIM-KS extends the language to include variables for elements of Σ and $\hat{K}_0$, extending term formation to cover these and is $PrAn + O\text{-}1\text{-}6 + C\text{-}1 + \sigma\text{-}\mu\text{-}1\text{-}2 + AX\text{-}MOD + BI_D\text{-}KS + BC\text{-}N\text{-}KS + \hat{K}_0 1\text{-}3 + AC\text{-}NN$. We also make use of the metamathematical notion of $\Sigma_\perp$ (§4.5).

We can obtain (as theorems) the following useful results: Generalised Open Data (Theorem 4.2.3.2), WC-N-KS (Theorem 4.2.3.3), and BI_D (Theorem 4.10.2)

6. Reduction to Analysis

§6.1. Roadmap to Analysis

In this chapter, we will demonstrate that $FIM\text{-}KS$, when placed under the constraints demanded by analysis, reduces down to a theory close to $FIM\text{-}AN$ (which contains our axioms of knowledge) given in the language of knowledge states.

Every branch of the conventional theory restricts the universe of choice sequences in some way to obtain the real number generators we use for analysis; our theory will be no different.

We will follow the path laid out in §1.5 and restrict ourselves to choice sequences of integers (the step from natural numbers to integers is achieved via a simple encoding) within the spread s_{tri} defined below.

$$s_{tri}(n) = \begin{cases} 0 \; iff \; \forall x_{0<x\leq(|n|-1)} \in N[2n(x)-1 \leq n(x+1) \leq 2n(x)+1] \\ 1 \; otherwise \end{cases}$$

where n is a finite sequence of integers

It is also from here that we enforce the schema OMR and $\exists KS\text{-}LS$ (see §4.6 for their justifications) which we have restated below to remind the reader.

OMR $\forall \underline{\mu} \exists \sigma[\forall \sigma' \subset \sigma[\sigma' \notin \Sigma_{SE}] \wedge \sigma(\underline{\mu}) \wedge \forall \sigma'[\forall \sigma'' \subseteq \sigma'[\sigma'' \notin \Sigma_{SE}] \wedge \sigma'(\underline{\mu}) \rightarrow \sigma' \subset \sigma]]$

$\exists KS\text{-}LS$ $\forall \mu \exists \nu \in M_{KSLS}[\mu = \nu]$

The reasons for this is that we will need OMR when constructing a KS-lawless sequence from any arbitrary sequence, the claim we proved using OMR in §4.6.

For the remainder of §6 all quantification over choice sequences will be restricted to the universe of choice sequences relativised to the spread s_{tri}. We also insist that the schema OMR is valid for these sequences, unless specifically indicated otherwise, in other words for all choice sequences μ in this chapter the following holds.

$$\exists\sigma[\forall\sigma' \subset \sigma[\sigma' \notin \Sigma_{SE}] \wedge \sigma(\mu) \wedge \forall\sigma'[\forall\sigma'' \subseteq \sigma'[\sigma'' \notin \Sigma_{SE}] \wedge \sigma'(\mu) \rightarrow \sigma' \subset \sigma]]$$

$$\wedge(Spread(s_{tri}))(\mu)$$

will be taken to hold. We will use the notation set up previously ($\mu \in M$) to denote quantification over choice sequences in general (if needed) for this chapter.

Finally, we introduce the assumption made to justify weak continuity (see §2.4) and restrict our quantification over knowledge states in certain instances in accordance with this. We interpret,

$\forall\underline{\mu}\exists\sigma$ to mean $\forall\underline{\mu}\exists\sigma_{|\sigma|\leq|\underline{\mu}|}$

$\forall\sigma\forall\underline{\mu}$ to mean $\forall\sigma\forall\underline{\mu}_{|\sigma|\leq|\underline{\mu}|}$

$\exists\sigma\forall\underline{\mu}$ to mean $\exists\sigma\forall\underline{\mu}_{|\sigma|\leq|\underline{\mu}|}$

This change gives the alternative interpretations of the following axioms.

σ-μ-1^* $\forall\underline{\mu}[A(\underline{\mu}) \rightarrow \exists\sigma[\sigma(\underline{\mu}) \wedge A'(\sigma)]]$, where A has no other choice sequence variables .

σ-μ-2^* $\forall\sigma[A'(\sigma) \rightarrow \forall\underline{\mu}[\sigma(\underline{\mu}) \rightarrow A(\underline{\mu})]]$, where A has no other free choice sequence variables.

BC-N-KS^* $\forall\underline{\mu}\exists\underline{x}[A(\underline{\mu}, \underline{x})] \rightarrow \exists\underline{\hat{e}}\forall\underline{\mu}\exists\sigma[\sigma(\underline{\mu}) \wedge A(\underline{\mu}, \underline{\hat{e}}(\sigma))]$, where A has no other free choice sequence variables.

$\hat{K}_0 1^*$ $\forall\underline{\mu}_{|\underline{\mu}|\geq ari(\hat{e})}\exists\sigma[\sigma(\underline{\mu}) \wedge \hat{e}(\sigma) \in N]$ (totality)

We will now write $\hat{e} \in \hat{K}_0$ to denote that $\hat{e}$ satisfied $\hat{K}_0 1^*$, $\hat{K}_0 2$ and $\hat{K}_0 3$.

Our goal is to show that our method of generating real numbers will be no different from the conventional notion explored in §1.5 and that our notion of continuity, when applied to the restrictions imposed by analysis, is the same as that in the conventional theory.

We shall continue in the following section to prove that for every continuous operation in analysis both the conventional notion of continuity and our new notion coincide.

The following section will outline the restriction of axioms used for FAN and UC to initial segments and how such restrictions are achieved in the language of knowledge states. We will also provide a comparison of our reduced system with $FIM\text{-}AN$, verifying that we do have all the components we require to obtain interesting analytic theorems such as FAN and UC.

The final section in this chapter will explore a notion $\hat{K}$ of inductively defined neighbourhood functions operating on knowledge states equivalent to the K of the conventional theory.

§6.2. Σ-Continuity and B-Continuity

In the conventional theory, we are interested in functions that are Baire Space continuous (B-continuous) for the purpose of analysis (Dummett, 1977 pp.79); this is reflected in the continuity axioms of the existing systems. We pause to remind the reader of the definitions of the conventional B-continuity and the Σ-continuity inherent in our system by stating both below.

> ψ is *B-continuous* iff $\forall\mu\exists x\forall\nu[\bar{\mu}(x) \subset \nu \rightarrow \psi(\mu) = \psi(\nu)]$.
>
> ψ is *Σ-continuous* iff $\forall\underline{\mu}_{|\underline{\mu}|=|\psi|}\exists\sigma[\sigma(\underline{\mu}) \wedge \forall\underline{\nu}_{|\underline{\nu}|=|\psi|}[\sigma(\underline{\nu}) \rightarrow \psi(\underline{\mu}) = \psi(\underline{\nu})]]$.

We can see that if ψ is B-continuous then this implies that ψ is Σ-continuous (the restriction of our finite information to initial segments) and also extensional ($\mu = \nu \rightarrow \psi(\mu) = \psi(\nu)$). In both cases, the converse is provably false as seen below.

> Take a function ψ such that $\psi(\mu) = \begin{cases} 0\ if\ \exists f[(Law(f))(\mu)] \\ 1\ otherwise \end{cases}$ (it is rendered decidable by the very fact that we enforce OMR). This function is Σ-continuous but it most certainly is **not** B-continuous.
>
> Take a function ψ such that $\psi(\mu) = \begin{cases} 0\ iff\ \mu = \nu \\ 1\ otherwise \end{cases}$, where ν is some pre-defined choice sequence. This function satisfies extensionality, however it most certainly is **not** B-continuous!

An important result asserted by the design of K_0 in the conventional theory is, ψ is B-continuous iff $\exists e \in K_0[e \sim \psi]$ (lemma 2.6.2).

We would also like to establish a similar relationship between Σ-continuous functions and $\hat{K}_0$ functions; $\exists \hat{e} \in \hat{K}_0[\hat{e} \sim \psi]$ iff ψ is Σ-continuous. Left to right is perfectly believable and trivially provable; however, right to left is a little trickier and will be given as a lemma.

Lemma 6.2.1

ψ is Σ-continuous $\rightarrow \exists \hat{e} \in \hat{K}_0[\hat{e} \sim \psi]$

Proof:

Given any ψ that is Σ-continuous define the predicate A as follows.

$A(\underline{\mu}, x) \iff \psi(\mu) = x$

Since ψ is a continuous operation defined for every $\underline{\mu}$ $(|\underline{\mu}| = |\psi|)$ we can assert that $\forall \underline{\mu} \exists x[A(\underline{\mu}, x)]$ where $\psi(\underline{\mu}) = x$.

Hence by BC-N-KS^* $\exists \hat{e} \in \hat{K}_0 \forall \underline{\mu} \exists \sigma[\sigma(\underline{\mu}) \wedge A(\underline{\mu}, \hat{e}(\sigma)]$.

Hence for our $\hat{e}$ we can rewrite this as $\forall \underline{\mu} \exists \sigma[\sigma(\underline{\mu}) \wedge \hat{e}(\sigma) = \psi(\underline{\mu})]$.

Thus by the definition of $\hat{e} \sim \psi$, this is equivalent to $\exists \hat{e} \in \hat{K}_0[\hat{e} \sim \psi]$ as required. ♠

With this final lemma we may now prove the following hard hitting reduction theorem.

Theorem 6.2.1 : The Reduction Theorem

If ψ is Σ-continuous and extensional then it is also B-continuous under the restrictions for analysis.

Proof:

For the sake of simplicity we will assume $|\psi| = 1$, however this proof easily generalises to a ψ of any arity via simple pairing and unpairing of sequences.

Assume ψ is both Σ-continuous and extensional.

By lemma 6.2.1 $\exists \hat{e} \in \hat{K}_0[\hat{e} \sim \psi]$, i.e. $\exists \hat{e} \in \hat{K}_0 \forall \mu \exists \sigma[\sigma(\mu) \wedge \hat{e}(\sigma) = \psi(\mu)]$.

Given our $\hat{e}$ and any arbitrary μ:

By $\hat{K}_0 1^*$ we can choose a $\sigma \in \Sigma$ such that $\sigma(\mu) \wedge \hat{e}(\sigma) \in N$.

By $\exists KS$-LS we can choose a $\nu \in M_{KSLS}$ such that $\mu = \nu$.

By $\hat{K}_0 1^*$ and the definition of M_{KSLS} we can choose a $\sigma' \in \Sigma_{SE}$ such that $\sigma'(\nu) \wedge \hat{e}(\sigma') \in N$.

By the extensionality of ψ, $\psi(\mu) = \psi(\nu)$.

$\hat{e}(\sigma) = \psi(\mu)$ by the definition of $\psi \sim \hat{e}$ and $\hat{K}_0 2$ and $\hat{K}_0 3$.

If $\hat{e}(\sigma) \neq \hat{e}(\sigma')$ then $\hat{e}(\sigma' = \Delta)$ by $\hat{K}_0 2$, but this cannot be the case by $\hat{K}_0 3$ so $\hat{e}(\sigma) = \hat{e}(\sigma')$.

By the fact that $\mu = \nu$ and lemma 4.6.1 we have $\sigma'(\mu)$.

Hence $\sigma'(\mu) \wedge \hat{e}(\sigma') \in N$.

By theorem 4.6.1 we can choose a $\sigma'' \in \Sigma_{IS}$ such that $\sigma' \subseteq \sigma''$ and $\sigma''(\mu)$.

By $\hat{K}_0 2$, $\hat{K}_0 3$ and the fact $\sigma' \subseteq \sigma''$ we have $\hat{e}(\sigma') = \hat{e}(\sigma'')$.

Hence, discharging our quantification over μ and setting $\sigma = \sigma''$ we have

$\forall\mu\exists\sigma \in \Sigma_{IS}[\sigma(\mu) \wedge \hat{e}(\sigma) \in N]$.

Define e as follows: $\forall n\forall\sigma \in \Sigma_{IS}[\sigma \sim n \rightarrow \hat{e}(\sigma) = e(n)]$.

By theorem 4.6.4 $\forall n\exists\sigma \in \Sigma_{IS}[\sigma \sim n]$.

So for any n we have a σ such that $e(n) = \hat{e}(\sigma)$ and since $\hat{e}$ is well defined then e must also be well defined.

By theorem 4.6.3 $\forall\sigma \in \Sigma_{IS}\exists n[\sigma \sim n]$.

Therefore $\forall\mu\exists\sigma \in \Sigma_{IS}[\sigma(\mu) \wedge \hat{e}(\sigma) \in N]$ is equivalent to $\forall\mu\exists n[n \subset \mu \wedge e(n) \in N]$. (1)

Given any μ such that $n \subset \mu$ then there is some $\sigma \in \Sigma_{IS}$ such that $e(n) = \hat{e}(\sigma) = \psi(\mu)$.

Hence (1) is equivalent to $\forall\mu\exists n[n \subset \mu \wedge e(n) = \psi(\mu)]$ which is equivalent to $e \sim \psi$ by definition.

For any n and m assume that $n \leq m$.

By theorem 4.6.4 we can choose σ and σ' in Σ_{IS} such that $n \sim \sigma$ and $m \sim \sigma'$.

$\sigma \subseteq \sigma'$ by lemma 4.6.2 and therefore $\hat{e}(\sigma) \preceq \hat{e}(\sigma')$ by $\hat{K}_0 2$.

This is equivalent to $e(n) \preceq e(m)$.

Discharging our assumptions and quantification over m and n this yields

$\forall n\forall m[n \leq m \rightarrow e(n) \preceq e(m)]$. (2)

Hence by (1) and (2) $e \in K_0$.

Hence $\exists e \in K_0[e \sim \psi]$.

Which is equivalent to ψ being Baire continuous by Lemma 2.6.2.

Discharging our hypothesis and quantification over ψ yields our desired result. ♠

This result essentially tells us that under the restrictions and conventions of analysis, the notion of Σ-continuity is equivalent to B-continuity, which is as good as saying 'the only information we require is initial segments'. While this notion has been argued throughout the conventional theory, I do not feel that anyone has actually performed such a reduction formally.

The following section will show that, under our requirement that a function is extensional and Σ-continuous, our generalised continuity axiom collapses down to the conventional one, and thus we shall have all the results we require to continue onto analysis as per §1.5 and §2.5.2.

§6.3. Final Axiom Reductions for Analysis

In this section, we establish our final reduction to analysis by showing that $BC\text{-}N\text{-}KS$ reduces to $BC\text{-}N$ for extensional A and hence provides a formal justification for $WC\text{-}N$. The theorem below shows that $BC\text{-}N$ is a special case of $BC\text{-}N\text{-}KS^*$ for extensional predicates.

Theorem 6.3.1

Given an extensional A then the schema $BC\text{-}N\text{-}KS^* \vdash BC\text{-}N$.

Proof:

Assume the hypothesis of $BC\text{-}N$ and that A is extensional.

By $BC\text{-}N\text{-}KS^*$, $\exists \hat{e} \in \hat{K}_0 \forall \mu \exists \sigma [\sigma(\mu) \wedge A(\mu, \hat{e}(\sigma))]$.

Given any μ,

We can choose a $\sigma \in \Sigma$ such that $\sigma(\mu) \wedge \hat{e}(\sigma) \in N$ by $\hat{K}_0 1^*$.

Therefore, by a similar argument given in theorem 6.2.1, we can also choose a $\sigma' \in \Sigma_{IS}$ such that $\sigma'(\mu) \wedge \hat{e}(\sigma') \in N$.

Discharging our quantification over μ and letting $\sigma' \equiv \sigma$ thus gives us

$\forall\mu\exists\sigma \in \Sigma_{IS}[\sigma(\mu) \wedge \hat{e}(\sigma) \in N]$.

Define e such that $\forall\sigma \in \Sigma_{IS}\forall n[\sigma \sim n \rightarrow \hat{e}(\sigma) = e(n)]$.

By a similar argument to theorem 6.2.1, e is total and $e \in K_0$.

Thus $\exists e \in K_0\forall\mu\exists n[\mu \subset n \wedge A(\mu, e(n))]$.

Discharging our assumptions we obtain BC-N as required. ♠

Based on theorem 4.10.2, the material presented in §1.5 and the comment at the end of §2.5.2, we now have all the results we need to obtain FAN and UC (the proofs are in §2.5.2).

The above theorem now allows us to state the following result.

Under the restrictions we have adopted for analysis, BI_D and BC-N (and hence WC-N and BI_M) are provable from the axioms of FIM-KS. Our spread restriction (s_{tri}) corresponds directly to the RNGs of §1.5 and all functions that we are interested in for analysis may be represented by K_0 functions.

Thus, under the restrictions for analysis, we have obtained every axiom schema of FIM save one – AC-NC. However, this schema is stronger than we need, and in actual fact, we have AC-NN, which, as seen in §3.4, is sufficient for our needs in analysis.

Of special note is that our notion of finite information (our knowledge states) once again become initial segments, under the restrictions for analysis in the given axioms. The same can said to be true of our knowledge axioms σ-μ-1 and σ-μ-2, the $\exists\underline{\nu}$ quantification vanishes in both via our assumption that no other choice sequence parameters may be included in the predicates, and this allows us to derive both open data and WC-N from these reduced forms.

Hence we have proven the following result.

Under the restrictions for analysis FIM-AN is derivable from FIM-KS.

Thus, any result provable in $FIM\text{-}AN$ is also provable in $FIM\text{-}KS$ under the restrictions for analysis, i.e. $FIM\text{-}KS$ is an extension of $FIM\text{-}AN$ under the restrictions for analysis.

§6.4. Inductively Defined Extended Neighbourhood Functions

This section aims to define a species of inductively defined functions that operate on knowledge states and are capable of representing **any** extensional Σ-continuous operation. The aim of this section is to demonstrate one particular way in which functions can make use of our knowledge states during function evaluation. It also provides solutions to the cycles presented at the end of §4.4.

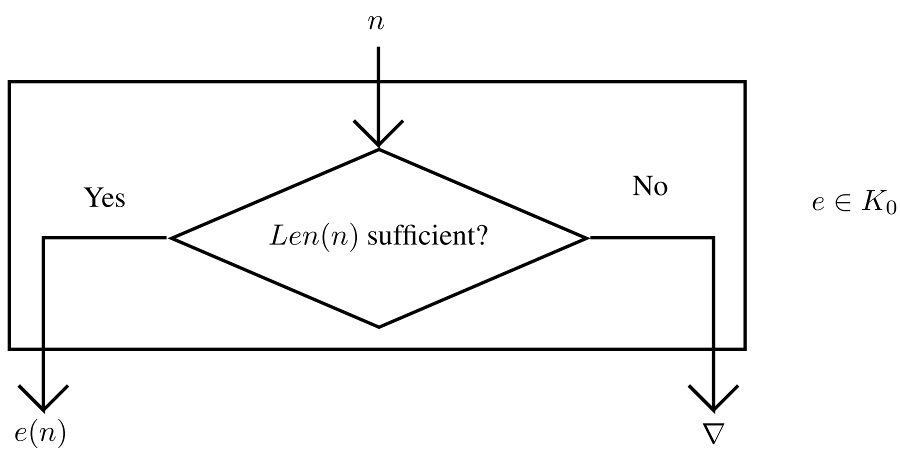
In §2.6, the idea behind the construction of K was that we start with the constant functions, and from those we construct functions that always give a natural number output if the length of the sequence input is > 0, and then, from these we construct functions that always give a natural number output if the length of the sequence input is > 1 and so on.

Given that the length of the input sequence is essentially used as a 'measure' of 'how much' information we have, we start with functions that require no information and produce functions that require more information. The first axiom $K1$ sets up these constant functions and the second axiom $K2$ provides the induction step used to construct more demanding functions from these. $K3$ simply states that K is the smallest possible set of these functions and is a combination of $K1$ and $K2$ with an additional clause at the end and some variable substitution. As such, we will construct this axiom last, once we have a satisfactory $\hat{K}1$ and $\hat{K}2$.

We will first look at how K_0 and K-functions are evaluated in the conventional theory before using these ideas to construct our species $\hat{K}$. We assume familiarity with λ notation during the remainder of this section.

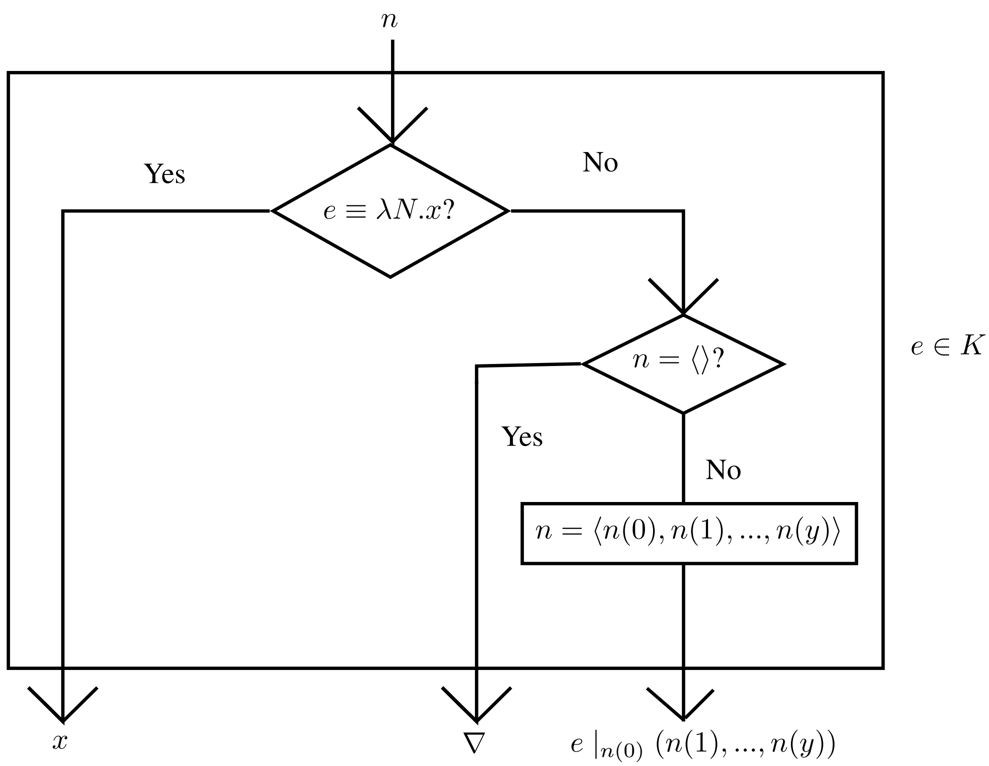
To begin with let us look at how a conventional neighbourhood function (**not necessarily** a K-function) would be calculated. The conventional theory dictates that we check if '[the finite sequence] n is long enough' (Troelstra and Dalen, 1988 §7.8 pp.211) in some way and if this is the case, then apply our neighbourhood function to the initial segment n.

Thus, we begin by following the method outlined in the conventional theory and assert that there is one step that occurs when we apply a *neighbourhood function* to a finite sequence. We perform some check to determine if there is enough information ($len(n)$ sufficient?) and then, if $len(n)$ is sufficient, output the evaluation of the continuous operation represented by the neighbourhood function ($e(n)$), or output the 'not enough information' value (∇) if the length of n is not sufficient. Figure 6 below gives a visual outline of this algorithm.



$Fig\ 6\ :\ K_0\ function$

However, when we move onto a K function, our method of evaluation is substantially altered. A K function does not have a length check as an explicit step, it is built in. Instead, it takes the sequence n, considers its first element and then produces another K function with that element 'built in' and applies this new neighbourhood function to the remaining elements (second onwards) of the sequence. Figure 7 below gives a visual outline of this algorithm.



$Fig\ 7\ :\ K\ function$

This concludes our exploration of evaluating the K_0 and K functions used in the conventional theory (and our own when restricted to analysis). We will now proceed to use these ideas to guide us in constructing a suitable notion for $\hat{K}$-functions.

In the conventional theory, we work with initial segments ($\sigma \in \Sigma_{IS}$), a concept that restricts us significantly. It only allows us to express one type of knowledge (SE) and it forces upon us a variable amount of 'redundant' information in each calculation. Our first refinement will be to extend neighbourhood functions to work with knowledge states in Σ_{SE}.

This change alters the way in which our function must look at the information it is given; in the conventional theory, the elements were always in order and there were no 'gaps', which allowed a K function to work using the first element at each step. In the extended theory, this is no longer the case as a $\hat{K}$ function cannot take the first piece of information it is given and cross section about it; this would lead to the same information given in a different order resulting in different outputs! We must also abandon the notion of checking for every element of lesser index than the ones we desire, as not all elements may be present. Instead, our $\hat{K}$-functions will have to be 'smarter', looking only for the elements required to evaluate the function and ignoring other information presented.

To do this, we shall use the notion of $\in$ defined earlier in §4.5 and also create a function that parses knowledge states and checks to see if the information we require is contained within; we shall call such functions *η-functions* and we say they are of type $\Sigma \to N^*$. For now, as we are only working with one type of atom (SE), we will define $\eta_{w,x}$ recursively as follows.

$$\eta_{w,x}(\sigma_\epsilon) = \begin{cases} y & \text{if } SE(w,x,y) \equiv \sigma_\epsilon \\ \nabla & \text{otherwise} \end{cases}$$

$$\eta_{w,x}(\sigma_1 \sqcap \sigma_2) = sup\{\eta_{w,x}(\sigma_1), \eta_{w,x}(\sigma_2)\}$$

$$\eta_{w,x}(\sigma_1 \sqcup \sigma_2) = inf\{\eta_{w,x}(\sigma_1), \eta_{w,x}(\sigma_2)\}$$

(We remind the reader that, when dealing with knowledge states about a single choice sequence, we will omit the w as it will always be 0).

An explanation as to why we sup with a $\sqcap$ and inf with a $\sqcup$ is provided below.

Assume we have the knowledge state $\sigma \equiv SE(0,0,1) \sqcap SE(0,0,2) \sqcap SE(0,1,3)$, a clearly contradictory knowledge state. Since our η-functions currently map $\Sigma_{SE} \to N^*$, they are responsible for 'detecting' such contradictions iff they are present in the information we are drawing elements from. η-functions work by exploring each 'branch' of a knowledge state and obtaining the values provided by the leaves of their branches; in the case of our example the leaves are the atomic knowledge states $SE(0,0,1)$, $SE(0,0,2)$ and $SE(0,1,3)$ giving us the following.

$$\eta_{0,0}(SE(0,0,1)) = 1$$

$$\eta_{0,0}(SE(0,0,2)) = 2$$

$$\eta_{0,0}(SE(0,1,3)) = \nabla$$

The reader may quickly notice that $sup\{1,2,\nabla\} = \Delta$, which is the exact value we want and this is why we use sup when dealing with $\sqcap$ cases.

If we had an $\hat{e}$ that outputs the second element of the sequence, then our η-function would be $\eta_{0,1}$ and,

$$\eta_{0,1}(SE(0,0,1)) = \nabla$$

$$\eta_{0,1}(SE(0,0,2)) = \nabla$$

$$\eta_{0,1}(SE(0,1,3)) = 3$$

$$sup\{\nabla,\nabla,3\} = 3$$

which is the answer we would expect. Hence, this idea holds very closely to the notions expressed in $\hat{K}_0 3$. We are not always obliged to output Δ, even if a knowledge state contains an obvious contradiction. We only forbid ourselves from outputting Δ when we are certain there are no contradictions in a knowledge state.

Our argument for the inf in $\sqcup$ cases forms a parallel. Assume we have the knowledge state $\sigma \equiv SE(0,0,1) \sqcup SE(0,0,2)$. This is not contradictory, it simply restricts the first element of a sequence to two possible values. We cannot be certain which of these values is the correct one and, thus, any function outputting the first element of a sequence acting on this would be expected to produce ∇. Our definition of how $\eta_{0,0}$ works gives us the following.

$$\eta_{0,0}(SE(0,0,1)) = 1$$

$$\eta_{0,0}(SE(0,0,2)) = 2$$

Thus, $inf(1,2) = \nabla$ which is the value we were expecting. The point here is that, if we have a two pieces of information, only one of which is true (the idea expressed by $\sqcup$), then unless both pieces of information agree we do not have enough information.

An example where $\sqcap$ and $\sqcup$ both occur is $\sigma \equiv (SE(0,0,1) \sqcup SE(0,0,2)) \sqcap SE(0,1,3)$.

$$\begin{aligned}
\eta_{0,0}(\sigma) &= \eta_{0,0}((SE(0,0,1) \sqcup SE(0,0,2)) \sqcap SE(0,1,3)) \\
&= sup(\eta_{0,0}(SE(0,0,1) \sqcup SE(0,0,2)), \eta_{0,0}(SE(0,1,3))) \\
&= sup(inf(\eta_{0,0}(SE(0,0,1)), \eta_{0,0}(SE(0,0,2))), \eta_{0,0}(SE(0,1,3))) \\
&= sup(inf(1,2), \nabla) \\
&= sup(\nabla, \nabla) \\
&= \nabla
\end{aligned}$$

$$\begin{aligned}
\eta_{0,1}(\sigma) &= \eta_{0,1}((SE(0,0,1) \sqcup SE(0,0,2)) \sqcap SE(0,1,3)) \\
&= sup(\eta_{0,1}(SE(0,0,1) \sqcup SE(0,0,2)), \eta_{0,1}(SE(0,1,3))) \\
&= sup(inf(\eta_{0,1}(SE(0,0,1)), \eta_{0,1}(SE(0,0,2))), \eta_{0,1}(SE(0,1,3))) \\
&= sup(inf(\nabla, \nabla), 3) \\
&= sup(\nabla, 3) \\
&= 3
\end{aligned}$$

Before we produce the next flowchart, a very important modification to the notion of cross section needs to be made. Previously, we were able to simply pass a number to a function and treat it like a parameter. However, in some cases we might end up having to pass one of the other elements of N^* (Δ or ∇) as a parameter. This causes us an unusual problem that is best shown with an example.

> Assume we have $\sigma \equiv SE(0,1) \sqcap SE(0,2)$, a clearly contradictory knowledge state, and an extended neighbourhood function that extracts the first element of a sequence (a function that uses the η-function $\eta_{0,0}$ to extract the first element of a sequence from a knowledge state and then cross sections to reduce down to a constant function). In this case we have the following.

$\eta_{0,0}(SE(0,1) \sqcap SE(0,2) = sup\{\eta_{0,0}(SE(0,0,1)), \eta_{0,0}(SE(0,0,2))\} = sup\{1,2\} = \Delta$

So, we end up having to cross section our extended neighbourhood function about Δ. The trouble is that this requires us to have a constant function that always outputs Δ which we want to avoid having in $\hat{K}$ (since such a function would violate $\hat{K}_0 1$). Similar examples exist for ∇, thus we need a 'work around' for this. A simple solution lies in modifying our definition of cross section to skip straight to an output if it is passed Δ or ∇ to avoid this issue.

Hence, for $\hat{e} \in \hat{K}$ we define $\hat{e} \downarrow_{x^*}$, where $x^* \in N^*$, as follows.

$$\hat{e} \downarrow_{x^*} = \begin{cases} \nabla & \text{iff } x^* = \nabla \\ \Delta & \text{iff } x^* = \Delta \\ \hat{e} \mid_{x^*} & \text{otherwise} \end{cases}$$

From now on we will use the shorthand notation $\hat{e} \downarrow_{\sigma x^*}$ to denote $\hat{e} \downarrow_{x^*} (\sigma)$ to keep our equations tidier.

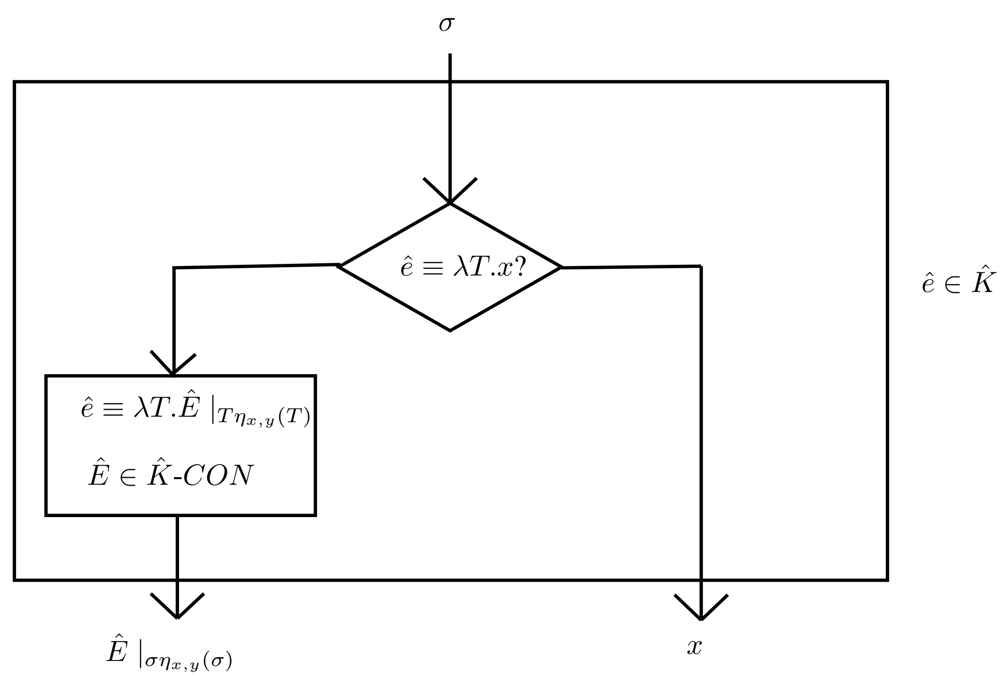
A notion tacitly used in K is a function of type $N \to (N^\omega \to N)$, a function that maps a natural number to a neighbourhood function. These are used in cross sectioning and we will need an analogue of this to properly construct our species $\hat{K}$. We call a function of type $N \to (\Sigma \to N^*)$ a $\hat{K}$*-constructor* and write this as $\hat{E} \in \hat{K}\text{-}CON$. The cross section of a $\hat{K}$ constructor by a natural number will always be a $\hat{K}$-function, i.e. $\forall x[E \mid_x \in K]$.

For $\hat{E} \in \hat{K}\text{-}CON$, we define $\hat{E} \downarrow_{x^*}$, where $x^* \in N^*$, as follows.

$$\hat{E} \downarrow_{x^*} = \begin{cases} \nabla & \text{iff } x^* = \nabla \\ \Delta & \text{iff } x^* = \Delta \\ \hat{E} \mid_{x^*} & \text{otherwise} \end{cases}$$

From now on we will use the shorthand notation $\hat{E} \downarrow_{\sigma x^*}$ to denote $\hat{E} \downarrow_{x^*} (\sigma)$ to keep our equations tidier.

With these changes, we may now outline our refined algorithm in figure 8 below.



$$Fig\ 8\ :\ \hat{K}\ function$$

The addition of law atoms simply requires the following refinement of our η-function.

$$\eta_{w,x}(\sigma_\epsilon) = \begin{cases} y & \text{if } SE(w,x,y) \equiv \sigma_\epsilon \\ f(x) & \text{if } law(w,f) \equiv \sigma_\epsilon \\ \nabla & \text{otherwise} \end{cases}$$

It is now a good time to investigate a case dealing with an η-function that we have currently avoided. What if we have some η-function that relies on some information provided earlier, for example, $\eta_{0,x}$, where x is the first element of sequence 0? Generally, given some η-function $\eta_{w,x}$ (a function of type $\Sigma \rightarrow N^*$), we may wish for the x component, w component or both components of this function to be calculated in some way from some natural number parameter. This function is not an η-function, as it has the wrong type $[N \rightarrow (\Sigma \rightarrow N^*)$ instead of $\Sigma \rightarrow N^*]$ but it is a notion we will require if we want to capture all the functions defined in the conventional theory; hence, we shall call any function of the type $N \to (\Sigma \to N^*)$ an *η-constructor.*

We now have a clear idea of what we require and may begin defining our species of $\hat{K}$-functions. We begin by defining our most basic kinds of $\hat{K}$-functions below.

$\hat{K}1$ $\exists x[\hat{e} \equiv \lambda T.x] \to \hat{e} \in \hat{K}$

This defines our constant functions. If we apply a knowledge state to the $\hat{e}$, and then use λ reduction, we obtain our output. For example, if we had $\hat{e} \equiv \lambda T.2$, then $(\hat{e})(SE(0,1) \sqcap SE(1,2)) \equiv 2$ and $(\hat{e})(SE(71,45)) \equiv 2$. This method gives us a clear relation between the intensional definition of a $\hat{K}$ function and its evaluation.

With our base case defined we may now construct the induction step for $\hat{K}$-functions, essentially allowing us to construct level functions requiring one element value from level 0 functions, functions requiring two element values from these functions requiring one element value, and so on. We define this induction step below.

$\hat{K}2*$ $(\hat{e} \equiv \lambda T.\hat{E} \downarrow_{T\eta(T)} \wedge \hat{E} \in \hat{K}\text{-}CON) \to \hat{e} \in \hat{K}$

Note that $\forall \hat{E} \in \hat{K}\text{-}CON\ \forall x[(\hat{E})(x) \in \hat{K}]$ still holds. These η-constructors and $\hat{K}$-constructors allow us to formulate $\hat{K}$-functions that either ask for a series of specific elements of a sequence or asks for one specific element and then other elements 'chosen' by that element in a functional way.

Finally, we need to say that $\hat{K}$ is the smallest such set of functions which we do below.

$$\hat{K}3* \ \forall \hat{e}[\exists x[(\hat{e} \equiv \lambda T.x)] \vee (\hat{e} \equiv \lambda T.\hat{E} \downarrow_{T\eta(T)} \wedge \forall x \in N[\hat{E} \downarrow_x \in Q]) \to \hat{e} \in Q]$$
$$\to \forall \hat{e}[\hat{e} \in \hat{K} \to \hat{e} \in Q]$$

To help better illustrate how these functions work, we refer the reader to the examples below.

Example 6.4.1 : $\hat{e}_1 \equiv \lambda T.((\lambda x \lambda T'.x)) \downarrow_{T\eta_0(T)}$ (A function that outputs the first element of a sequence).

Knowledge State : $SE(0,1)$

Expected Output : 1

$$\hat{e}_1(SE(0,1)) \equiv (\lambda T.(\lambda x \lambda T'.x) \downarrow_{T\eta_0(T)})(SE(0,1))$$

$$\equiv (\lambda x \lambda T'.x) \downarrow_{SE(0,1)\eta_0(SE(0,1))}$$

$$\eta_0(SE(0,1)) = 1$$

$$\hat{e}_1(SE(0,1)) \equiv (\lambda x \lambda T'.x) \downarrow_{SE(0,1)1}$$

$$\equiv (\lambda T'.1) \downarrow_{SE(0,1)}$$

$$\equiv 1$$

Knowledge State : $SE(1,1)$

Expected Output : ∇

$$\hat{e}_1(SE(1,1)) \equiv (\lambda T.(\lambda x \lambda T'.x) \downarrow_{T\eta_0(T)})(SE(1,1))$$

$$\equiv (\lambda x \lambda T'.x) \downarrow_{SE(0,1)\eta_0(SE(1,1))}$$

$$\eta_0(SE(1,1)) = \nabla$$

$$\hat{e}_1(SE(1,1)) \equiv (\lambda x \lambda T'.x) \downarrow_{SE(1,1)\nabla}$$

$$\equiv \nabla$$

Knowledge State : $SE(0,1) \sqcap SE(1,1)$

Expected Output : 1

$$\hat{e}_1(SE(0,1) \sqcap SE(1,1)) \equiv (\lambda T.(\lambda x \lambda T'.x) \downarrow_{T\eta_0(T)})(SE(0,1) \sqcap SE(1,1))$$

$$\equiv (\lambda x \lambda T'.x) \downarrow_{(SE(0,1)\sqcap SE(1,1))\eta_0(SE(0,1)\sqcap SE(1,1))}$$

$$\eta_0(SE(0,1) \sqcap SE(1,1)) = sup(\eta_0(SE(0,1)), \eta_0(SE(1,1)))$$

$$\eta_0(SE(0,1)) = 1$$

$$\eta_0(SE(1,1)) = \nabla$$

$$\eta_0(SE(0,1) \sqcap SE(1,1)) = sup(1, \nabla)$$

$$= 1$$

$$\hat{e}_1(SE(0,1) \sqcap SE(1,1)) \equiv (\lambda x \lambda T'.x) \downarrow_{(SE(0,1) \sqcap SE(1,1))1}$$

$$\equiv (\lambda T'.1) \downarrow_{(SE(0,1) \sqcap SE(1,1))}$$

$$\equiv 1$$

Knowledge State : $SE(0,1) \sqcap SE(0,2)$

Expected Output : Δ

$$\hat{e}_1(SE(0,1) \sqcap SE(0,2)) \equiv (\lambda T.(\lambda x \lambda T'.x) \downarrow_{T\eta_0(T)})(SE(0,1) \sqcap SE(0,2))$$

$$\equiv (\lambda x \lambda T'.x) \downarrow_{(SE(0,1) \sqcap SE(0,2))\eta_0(SE(0,1) \sqcap SE(0,2))}$$

$$\eta_0(SE(0,1) \sqcap SE(0,2)) = sup(\eta_0(SE(0,1)), \eta_0(SE(0,2)))$$

$$\eta_0(SE(0,1)) = 1$$

$$\eta_0(SE(0,2)) = 2$$

$$\eta_0(SE(0,1) \sqcap SE(0,2)) = sup(1,2)$$

$$= \Delta$$

$$\hat{e}_1(SE(0,1) \sqcap SE(0,2)) \equiv (\lambda x \lambda T'.x) \downarrow_{(SE(0,1) \sqcap SE(0,2))\Delta}$$

$$\equiv \Delta$$

Knowledge State : $law(\lambda x.x)$

Expected Output : 0

$$\hat{e}_1(law(\lambda x.x)) \equiv (\lambda T.(\lambda x \lambda T'.x) \downarrow_{T\eta_0(T)})(law(\lambda x.x))$$

$$\equiv (\lambda x \lambda T'.x) \downarrow_{law(\lambda x.x)\eta_0(law(\lambda x.x))}$$

$$\eta_0(law(\lambda x.x)) = 0$$

$$\hat{e}_1(law(\lambda x.x)) \equiv (\lambda x \lambda T'.x) \downarrow_{law(\lambda x.x)0}$$

$$\equiv (\lambda T'.0) \downarrow_{law(\lambda x.x)}$$

$$\equiv 0$$

Knowledge State : $law(\lambda x.x) \sqcap SE(0,1)$

Expected Output : Δ

$$\hat{e}_1(law(\lambda x.x) \sqcap SE(0,1)) \equiv (\lambda T.(\lambda x \lambda T'.x) \downarrow_{T\eta_0(T)})(law(\lambda x.x) \sqcap SE(0,1))$$

$$\equiv (\lambda x \lambda T'.x) \downarrow_{(law(\lambda x.x) \sqcap SE(0,1))\eta_0(law(\lambda x.x) \sqcap SE(0,1))}$$

$$\eta_0(law(\lambda x.x) \sqcap SE(0,1)) = sup(\eta_0(law(\lambda x.x)), \eta_0(SE(0,1)))$$

$$\eta_0(law(\lambda x.x)) = 0$$

$$\eta_0(SE(0,1)) = 1$$

$$\eta_0(law(\lambda x.x) \sqcap SE(0,1)) = sup(0,1)$$

$$= \Delta$$

$$\hat{e}_1(law(\lambda x.x) \sqcap SE(0,1)) \equiv (\lambda x \lambda T'.x) \downarrow_{(law(\lambda x.x) \sqcap SE(0,1))\Delta}$$

$$\equiv \Delta$$

Example 6.4.2: $\hat{e}_2 \equiv \lambda T.(\lambda x.\lambda T'.(\lambda x'.\lambda T''.x{+}x') \downarrow_{T'\eta_3(T')}) \downarrow_{T\eta_0(T)}$ (A function that outputs the sum of the first and fourth elements of a sequence).

Knowledge State : $SE(0,1)$

Expected Output : ∇

$$\hat{e}_2(SE(0,1)) \equiv (\lambda T.(\lambda x.\lambda T'.(\lambda x'.\lambda T''.x + x') \downarrow_{T'\eta_3(T')}) \downarrow_{T\eta_0(T)})(SE(0,1))$$

$$\equiv (\lambda x.\lambda T'.(\lambda x'.\lambda T''.x + x') \downarrow_{T'\eta_3(T')}) \downarrow_{SE(0,1)\eta_0(SE(0,1))}$$

$$\eta_0(SE(0,1)) = 1$$

$$\hat{e}_2(SE(0,1)) \equiv (\lambda x.\lambda T'.(\lambda x'.\lambda T''.x + x') \downarrow_{T'\eta_3(T')}) \downarrow_{SE(0,1)1}$$

$$\equiv (\lambda T'.(\lambda x'.\lambda T''.1 + x') \downarrow_{T'\eta_3(T')}) \downarrow_{SE(0,1)}$$

$$\equiv (\lambda x'.\lambda T''.1 + x') \downarrow_{SE(0,1)\eta_3(SE(0,1))}$$

$$\eta_3(SE(0,1)) = \nabla$$

$$\hat{e}_2(SE(0,1)) \equiv (\lambda x'.\lambda T''.1 + x') \downarrow_{SE(0,1)\nabla}$$

$$\equiv \nabla$$

Knowledge State : $SE(3,1)$

Expected Output : ∇

$$\hat{e}_2(SE(3,1)) \equiv (\lambda T.(\lambda x.\lambda T'.(\lambda x'.\lambda T''.x + x') \downarrow_{T'\eta_3(T')}) \downarrow_{T\eta_0(T)})(SE(3,1))$$

$$\equiv (\lambda x.\lambda T'.(\lambda x'.\lambda T''.x + x') \downarrow_{T'\eta_3(T')}) \downarrow_{SE(3,1)\eta_0(SE(3,1))}$$

$$\eta_0(SE(3,1)) = \nabla$$

$$\hat{e}_2(SE(3,1)) \equiv (\lambda x.\lambda T'.(\lambda x'.\lambda T''.x + x') \downarrow_{T'\eta_3(T')}) \downarrow_{SE(3,1)\nabla}$$

$$\equiv \nabla$$

Knowledge State : $SE(0,1) \sqcap SE(3,2)$

Expected Output : 3

$$\hat{e}_2(SE(0,1) \sqcap SE(3,2)) \equiv (\lambda T.(\lambda x.\lambda T'.(\lambda x'.\lambda T''.x + x') \downarrow_{T'\eta_3(T')}) \downarrow_{T\eta_0(T)})(SE(0,1) \sqcap SE(3,2))$$

$$\equiv (\lambda x.\lambda T'.(\lambda x'.\lambda T''.x + x') \downarrow_{T'\eta_3(T')}) \downarrow_{(SE(0,1)\sqcap SE(3,2))\eta_0(SE(0,1\sqcap SE(3,2)))}$$

$$\eta_0(SE(0,1) \sqcap SE(3,2)) = sup(\eta_0(SE(0,1)), \eta_0(SE(3,2)))$$

$$\eta_0(SE(0,1)) = 1$$

$$\eta_0(SE(3,2)) = \nabla$$

$$\eta_0(SE(0,1) \sqcap SE(3,2)) = sup(1, \nabla)$$

$$= 1$$

$$\hat{e}_2(SE(0,1) \sqcap SE(3,2)) \equiv (\lambda x.\lambda T'.(\lambda x'.\lambda T''.x + x') \downarrow_{T'\eta_3(T')}) \downarrow_{(SE(0,1)\sqcap SE(3,2))1}$$

$$\equiv (\lambda T'.(\lambda x'.\lambda T''.1 + x') \downarrow_{T'\eta_3(T')}) \downarrow_{SE(0,1)\sqcap SE(3,2)}$$

$$\equiv (\lambda x'.\lambda T''.1 + x') \downarrow_{(SE(0,1)\sqcap SE(3,2))\eta_3(SE(0,1)\sqcap SE(3,2))}$$

$$\eta_3(SE(0,1) \sqcap SE(3,2)) = sup(\eta_3(SE(0,1)), \eta_3(SE(3,2)))$$

$$\eta_3(SE(0,1)) = \nabla$$

$$\eta_3(SE(3,2)) = 2$$

$$\eta_3(SE(0,1) \sqcap SE(3,2)) = sup(\nabla, 2)$$

$$= 2$$

$$\hat{e}_2(SE(0,1)) \equiv (\lambda x'.\lambda T''.1 + x') \downarrow_{(SE(0,1) \sqcap SE(3,2))2}$$

$$\equiv (\lambda x'.\lambda T''.1 + x') \downarrow_{(SE(0,1) \sqcap SE(3,2))2}$$

$$\equiv (\lambda T''.1 + 2) \downarrow_{(SE(0,1) \sqcap SE(3,2))}$$

$$\equiv 1 + 2$$

$$\equiv 3$$

Knowledge State : $SE(0,1) \sqcup SE(3,2)$

Expected Output : ∇

$$\hat{e}_2(SE(0,1) \sqcup SE(3,2)) \equiv (\lambda T.(\lambda x.\lambda T'.(\lambda x'.\lambda T''.x + x') \downarrow_{T'\eta_3(T')}) \downarrow_{T\eta_0(T)})(SE(0,1) \sqcup SE(3,2))$$

$$\equiv (\lambda x.\lambda T'.(\lambda x'.\lambda T''.x + x') \downarrow_{T'\eta_3(T')}) \downarrow_{(SE(0,1) \sqcup SE(3,2))\eta_0(SE(0,1 \sqcup SE(3,2)))}$$

$$\eta_0(SE(0,1) \sqcup SE(3,2)) = inf(\eta_0(SE(0,1)), \eta_0(SE(3,2)))$$

$$\eta_0(SE(0,1)) = 1$$

$$\eta_0(SE(3,2)) = \nabla$$

$$\eta_0(SE(0,1) \sqcup SE(3,2)) = inf(1, \nabla)$$

$$= \nabla$$

$$\hat{e}_2(SE(0,1) \sqcup SE(3,2)) \equiv (\lambda x.\lambda T'.(\lambda x'.\lambda T''.x + x') \downarrow_{T'\eta_3(T')}) \downarrow_{(SE(0,1) \sqcup SE(3,2))\nabla}$$

$$\equiv \nabla$$

Knowledge State : $\sigma \equiv SE(0,1) \sqcap SE(3,2) \sqcap law(\lambda x.x + 1)$

Expected Output : Δ

$$\hat{e}_2(\sigma) \equiv (\lambda T.(\lambda x.\lambda T'.(\lambda x'.\lambda T''.x + x') \downarrow_{T'\eta_3(T')}) \downarrow_{T\eta_0(T)})(\sigma)$$

$$
\begin{aligned}
&\equiv (\lambda x.\lambda T'.(\lambda x'.\lambda T''.x + x') \downarrow_{T' \eta_3(T')}) \downarrow_{\sigma \eta_0(\sigma)} \\
\eta_0(\sigma) &= sup(\eta_0(SE(0,1)), \eta_0(SE(3,2) \sqcap law(\lambda x.x+1))) \\
&= sup(\eta_0(SE(0,1)), sup(\eta_0(SE(3,2)), \eta_0(law(\lambda x.x+1)))) \\
\eta_0(SE(0,1)) &= 1 \\
\eta_0(SE(3,2)) &= \nabla \\
\eta_0(law(\lambda x.x+1)) &= 1 \\
\eta_0(\sigma) &= sup(1, sup(\nabla, 1)) \\
&= sup(1,1) \\
&= 1 \\
\hat{e}_2(SE(0,1) \sqcap SE(3,2)) &\equiv (\lambda x.\lambda T'.(\lambda x'.\lambda T''.x + x') \downarrow_{T' \eta_3(T')}) \downarrow_{(SE(0,1) \sqcap SE(3,2))1} \\
&\equiv (\lambda T'.(\lambda x'.\lambda T''.1 + x') \downarrow_{T' \eta_3(T')}) \downarrow_{SE(0,1) \sqcap SE(3,2)} \\
&\equiv (\lambda x'.\lambda T''.1 + x') \downarrow_{(SE(0,1) \sqcap SE(3,2)) \eta_3(SE(0,1) \sqcap SE(3,2))} \\
\eta_3(\sigma) &= sup(\eta_3(SE(0,1)), \eta_3(SE(3,2) \sqcap law(\lambda x.x+1))) \\
&= sup(\eta_3(SE(0,1)), sup(\eta_3(SE(3,2)), \eta_3(law(\lambda x.x+1)))) \\
\eta_0(SE(0,1)) &= \nabla \\
\eta_0(SE(3,2)) &= 2 \\
\eta_0(law(\lambda x.x+1)) &= 4 \\
\eta_0(\sigma) &= sup(\nabla, sup(2,4)) \\
&= sup(1, \Delta) \\
&= \Delta \\
\hat{e}_2(SE(0,1)) &\equiv (\lambda x'.\lambda T''.1 + x') \downarrow_{\sigma \Delta} \\
&\equiv \Delta
\end{aligned}
$$

We now extend ourselves yet further and admit the third kind of information to be worked upon by the extended neighbourhood functions, LR atoms. We pause here briefly to remind the reader that we are only considering extensional operations in this section. As such, all functions mentioned are assumed to satisfy extensionality, including the R in LR atoms.

Looking closely at LR atoms (see §4.4), note the similarity between the relation R and the f in our law atom **for the cases where R is extensional**.

> The f in law atoms is passed an index value, and, from there, applies a function to it to obtain a natural number, which is the desired element of a choice sequence.
>
> Any extensional relation R should work in a similar way; we pass it an index and it then uses information from the tuple of choice sequences $\underline{\nu}$ to find the appropriate value of μ.

The question of 'how does it do this?' raises an interesting thought, since, if we are applying a function to a choice sequence (or tuple of), we must use a neighbourhood function. We essentially pass something a numerical variable and it then provides us with a neighbourhood function that calculates the value of μ we desire. The fact that law relations are 'relations that allow us to generate the elements of one sequence from information about a tuple of other sequences' (§4.4), and the parametrisation comment in the paragraph above, ensures us that the Rs in LR atoms are of type $N \to (\Sigma \to N^*)$. In other words, they meet our requirements to be a $\hat{K}$-constructor.

Let us assume that in the course of evaluating some $\hat{e}$ we call upon an η-function that uses a LR atom, formally written below.

$$\eta_{w,x}(LR(R, w, y_1, .., y_z))$$

To evaluate this, we first pass R the index of the element we desire (x) and then we apply it to some knowledge state as follows.

$$R\mid_x(\sigma)$$

The question is, what knowledge do we apply this to? Given the way η-functions currently work, the only information it 'holds' is about the knowledge state it is evaluating and $R\mid_x (LR(R,w,y_1,...y_z))$ would be almost guaranteed to output ∇. What we would like is for our η-function to be able to pass $R\mid_x$ the sum total of our knowledge (the original σ of which $LR(R,w,y_1,...,y_z)$ is an atom). To do this, we need to slightly modify our notion of an η-function. Thus we make the following revisions to our definitions of η-functions and η-constructors.

$$\eta_{w,x}(\sigma_1 \sqcap \sigma_2, \sigma_3) = sup(\eta_{w,x}(\sigma_1,\sigma_3), \eta_{w,x}(\sigma_2,\sigma_3))$$

$$\eta_{w,x}(\sigma_1 \sqcup \sigma_2, \sigma_3) = inf(\eta_{w,x}(\sigma_1,\sigma_3), \eta_{w,x}(\sigma_2,\sigma_3))$$

$$\eta_{w,x}(\sigma_\epsilon, \sigma) = \begin{cases} y & \text{if } SE(w,x,y) \equiv \sigma_\epsilon \\ f(x) & \text{if } law(w,f) \equiv \sigma_\epsilon \\ (Rx)(\sigma) & \text{if } LR(R,w,\underline{w}) \equiv \sigma_\epsilon \\ \nabla & \text{otherwise} \end{cases}$$

η-functions are now of type $\Sigma \times \Sigma \to N^*$.

An η-constructor is any function of the type $N^x \to (\Sigma \times \Sigma \to N^*)$.

The η-function above is not quite right, as it is possible to obtain a cycle where evaluation never ends. Assume we have the $\hat{K}$ function,

$\hat{e} \equiv \lambda T.[\lambda x_1.\lambda T'.x_1] \downarrow_{T\eta_{0,0}(T,T)}$ (the function outputting the first element of the first sequence)

and the knowledge state,

$\sigma \equiv LR(R_1, 0, 1) \sqcap LR(R_2, 1, 0)$

where $R_1 \equiv \lambda x.\lambda T.(\lambda x'.\lambda T'.x') \downarrow_{T\eta_{1,x}(T,T)}$ and $R_2 \equiv \lambda x.\lambda T.(\lambda x'.\lambda T'.x') \downarrow_{T\eta_{0,x}(T,T)}$ (a pair of symmetrical relationships).

Then we have the following.

$$\hat{e}(\sigma) \equiv (\lambda T.[\lambda x_1.\lambda T'.x_1] \downarrow_{T\eta_{0,0}(T,T)})(\sigma)$$

$$\equiv \lambda T.[\lambda x_1.\lambda T'.x_1] \downarrow_{\sigma\eta_{0,0}(\sigma,\sigma)}$$

$$\eta_{0,0}(\sigma,\sigma) = \eta_{0,0}(LR(R_1,0,1) \sqcap LR(R_2,1,0),\sigma)$$

$$= sup(\eta_{0,0}(LR(R_1,0,1)\sigma),\eta_{0,0}(LR(R_2,1,0),\sigma))$$

$$\eta_{0,0}(LR(R_1,0,1),\sigma) = (\lambda T.(\lambda x'.\lambda T'.x') \downarrow_{T\eta_{1,0}(T,T)})(\sigma)$$

$$= (\lambda T.(\lambda x'.\lambda T'.x') \downarrow_{\sigma\eta_{1,0}(\sigma,\sigma)})$$

$$\eta_{1,0}(\sigma,\sigma) = \eta_{1,0}(LR(R_1,0,1) \sqcap LR(R_2,1,0),\sigma)$$

$$= sup(\eta_{1,0}(LR(R_1,0,1),\sigma),\eta_{1,0}(LR(R_2,1,0),\sigma))$$

$$\eta_{1,0}(LR(R_1,0,1),\sigma) = \nabla$$

$$\eta_{1,0}(LR(R_2,1,0),\sigma) = (\lambda T.(\lambda x'.\lambda T'.x') \downarrow_{T\eta_{0,0}(T,T)})(\sigma)$$

$$= (\lambda T.(\lambda x'.\lambda T'.x') \downarrow_{\sigma\eta_{0,0}(\sigma,\sigma)})$$

$$\eta_{0,0}(\sigma,\sigma) = ...$$

We were already trying to evaluate $\eta_{0,0}(\sigma,\sigma)$, hence, we have an evaluation that will never end, i.e. a cycle.

Thus, in the interest of avoiding cycles, we modify our η-function to prevent this issue as follows:

$$\eta_{w,x}(\sigma_\epsilon,\sigma) = \begin{cases} y & if\ SE(w,x,y) \equiv \sigma_\epsilon \\ f(x) & if\ law(w,f) \equiv \sigma_\epsilon \\ (Rx)(\sigma - \sigma_\epsilon) & if\ LR(R,w,\underline{w}) \equiv \sigma_\epsilon \\ \nabla & otherwise \end{cases}$$

where $\sigma - \sigma_\epsilon$ is defined recursively as follows,

$$\sigma - \sigma_\epsilon \equiv \begin{cases} (\sigma_1 - \sigma_\epsilon) \sqcap (\sigma_2 - \sigma_\epsilon) & \iff \sigma \equiv \sigma_1 \sqcap \sigma_2 \\ (\sigma_1 - \sigma_\epsilon) \sqcup (\sigma_2 - \sigma_\epsilon) & \iff \sigma \equiv \sigma_1 \sqcup \sigma_2 \\ \sigma_\emptyset & \iff \sigma \equiv \sigma_\epsilon \vee \sigma \equiv \sigma_\emptyset \end{cases}$$

Now, the example given above evaluates as follows:

$$
\begin{aligned}
\hat{e}(\sigma) &\equiv (\lambda T.[\lambda x_1.\lambda T'.x_1] \downarrow_{T\eta_{0,0}(T,T)})(\sigma) \\
&\equiv \lambda T.[\lambda x_1.\lambda T'.x_1] \downarrow_{\sigma\eta_{0,0}(\sigma,\sigma)} \\
\eta_{0,0}(\sigma,\sigma) &= \eta_{0,0}(LR(R_1,0,1) \sqcap LR(R_2,1,0),\sigma) \\
&= sup(\eta_{0,0}(LR(R_1,0,1)\sigma),\eta_{0,0}(LR(R_2,1,0),\sigma)) \\
\eta_{0,0}(LR(R_1,0,1),\sigma) &= (\lambda T.(\lambda x'.\lambda T'.x') \downarrow_{T\eta_{1,0}(T,T)})(LR(R_2,1,0)) \\
&= \lambda T.(\lambda x'.\lambda T'.x') \downarrow_{LR(R_2,1,0)\eta_{1,0}(LR(R_2,1,0),LR(R_2,1,0))} \\
\eta_{1,0}(LR(R_2,1,0),LR(R_2,1,0)) &= (\lambda T.(\lambda x'.\lambda T'.x') \downarrow_{T\eta_{0,0}(T,T)})(\sigma_\emptyset) \\
&= \lambda T.(\lambda x'.\lambda T'.x') \downarrow_{\sigma_\emptyset\eta_{0,0}(\sigma_\emptyset,\sigma_\emptyset)} \\
\eta_{0,0}(\sigma_\emptyset,\sigma_\emptyset) &= \nabla \\
\eta_{1,0}(LR(R_2,1,0),LR(R_2,1,0)) &= (\lambda T.(\lambda x'.\lambda T'.x') \downarrow_{\sigma\nabla)}) \\
&= \nabla \\
\eta_{0,0}(LR(R_1,0,1),\sigma) &= \lambda T.(\lambda x'.\lambda T'.x') \downarrow_{LR(R_2,1,0)\nabla} \\
&= \nabla \\
\eta_{0,0}(LR(R_2,1,0),\sigma) &= \nabla \\
\eta_{0,0}(\sigma,\sigma) &= sup(\nabla,\nabla) \\
&= \nabla \\
\hat{e}(\sigma) &\equiv \lambda T.[\lambda x_1.\lambda T'.x_1] \downarrow_{\sigma\nabla} \\
&\equiv \nabla
\end{aligned}
$$

This output seems far more reasonable than an infinite loop; it is certainly what a human evaluating this function would conclude.

We offer our final formulation of $\hat{K}$ below.

$\hat{K}1$ $\exists x[\hat{e} \equiv \lambda T.x] \rightarrow \hat{e} \in \hat{K}$

$\hat{K}2$ $(\hat{e} \equiv \lambda T.\hat{E} \downarrow_{T\eta(T,T)} \wedge \hat{E} \in \hat{K}\text{-}CON) \to \hat{e} \in \hat{K}$

$\hat{K}3$ $\forall \hat{e}[\exists x[(\hat{e} \equiv \lambda T.x)] \vee (\hat{e} \equiv \lambda T.\hat{E} \downarrow_{T\eta(T,T)} \wedge \forall x \in N[\hat{E} \downarrow_x \in Q]) \to \hat{e} \in Q] \to \forall \hat{e}[\hat{e} \in \hat{K} \to \hat{e} \in Q]$

We pause to raise a final critical point before we conclude our construction of $\hat{K}$; what of fan, *spread* and SR knowledge states? For this we direct the reader to §7.3.7, and leave the matter to rest for now. This concludes our construction of $\hat{K}$ and we will now prove the following vital theorem.

Theorem 6.4.1

$\hat{K} \subset \hat{K}_0$

Proof:

We begin as we did in the with the proof in the conventional theory, setting $Q = \hat{K}_0$ in $\hat{K}3$.

There are two types of function we must consider – those of the form given by $\hat{K}1$ and those of the form given in $\hat{K}2$.

Assume $\hat{e}$ is of the form given in $\hat{K}1$.

$\exists x \forall \sigma[\hat{e}(\sigma) = x]$ so $\hat{K}_0 1$ and $\hat{K}_0 2$ will always be satisfied.

$\lambda T.\Delta$ is not a function of the form given by $\hat{K}1$, so functions of this type can never output Δ, and hence $\hat{K}_0 3$ is also always satisfied.

Hence, any functions of the form given in $\hat{K}1$ are in $\hat{K}_0$.

Let us now assume we have an $\hat{e}$ of the form given in $\hat{K}2$, i.e. $\hat{e} \equiv \lambda T.(\hat{E}) \downarrow_{T\eta_{w,x}(T,T)}$, where $\forall x \in N[E \mid_x \in \hat{K}_0]$.

Take any $\underline{\mu}$ such that $|\underline{\mu}| \geq ari(\hat{e})$, and the knowledge state $SE(w,x,y)$ such that $(SE(w,x,y))(\underline{\mu})$.

Then our $\hat{e}$ would reduce to $(\hat{E})\downarrow_{SE(w,x,y)\eta_{w,x}(SE(w,x,y),SE(w,x,y))} \equiv \hat{E}\mid_y (SE(w,x,y))$.

Since $\hat{E}\mid_y \in \hat{K}_0$ then, for our given $\underline{\mu}$, $\exists\sigma[\sigma(\mu) \wedge \hat{E}\mid_y(\sigma) \in N]$ by $\hat{K}_0 1$.

Hence, $\hat{e}(\sigma \sqcap SE(w,x,y))$ will also output a natural number.

Discharging our quantification over $\underline{\mu}$,means that $\forall\underline{\mu}_{|\underline{\mu}|\geq ari(\hat{e})}\exists\sigma'[\sigma'(\underline{\mu}) \wedge \hat{e}(\sigma') \in N]$,

and hence $\hat{e}$ satisfies $\hat{K}_0 1$.

Given any σ and σ' assume that $\sigma \subseteq \sigma'$.

If $\hat{e}(\sigma') = \nabla$ then, by the definition of $\subseteq$, $\hat{e}(\sigma) = \nabla$.

If $\hat{e}(\sigma') \in N$ then define $y_0 = \eta_{w,x}(\sigma')$, then $\exists x[\hat{E}\downarrow_{y_0}(\sigma') = \hat{e}(\sigma')]$.

$\hat{E}\downarrow_{y_0}(\sigma) \preceq \hat{E}\downarrow_{y_0}(\sigma')$, and hence $\hat{e}(\sigma) \preceq \hat{e}(\sigma')$, if

$\eta_{w,x}(\sigma') = \eta_{w,x}(\sigma) \vee \eta_{w,x}(\sigma') = \eta_{w,x}(\sigma'')$.

Assume that this isn't the case and that $\eta_{w,x}(\sigma) = y_1$ and $\eta_{w,x}(\sigma'') = y_2$, then $\eta_{w,x}(\sigma') = sup(y_1, y_2)$, which must equal y_0. The only possible way this can occur is if $y_1 = y_0$ or $y_2 = y_0$ or $y_1 = y_0 = y_2$; in each of these cases the condition above is satisfied.

Hence, $\hat{e}(\sigma) \preceq \hat{e}(\sigma')$.

If $\hat{e}(\sigma') = \Delta$, then, by the definition of $\preceq$, $\hat{e}(\sigma) = \Delta$.

Discharging our hypothesis and quantification over σ and σ' yields $\forall\sigma\forall\sigma'[\sigma \subseteq \sigma' \to \hat{e}(\sigma) \preceq \hat{e}(\sigma')]]$, as required for $\hat{e}$ to satisfy $\hat{K}_0 2$.

Given any σ, assume that $\exists\underline{\mu}$ such that $\sigma(\underline{\mu})$.

Then $\forall x \in n[\hat{E}\mid_x(\sigma) \prec \Delta]$, since $\forall x \in N[\hat{E}\mid_x \in \hat{K}_0]$.

Thus the only way for $\hat{e}(\sigma) = \Delta$ is if $\eta_{w,x}(\sigma) = \Delta$.

The only way for this to occur is if $\sigma \to \exists y \exists z[\underline{\mu}_w(x) = y \wedge \underline{\mu}_w(x) = z]$, but this contradicts the statement that $\sigma(\underline{\mu})$!

Thus $\hat{e}(\sigma) \prec \Delta$.

Discharging our quantifier and hypothesis, we have that $\forall\sigma[\exists\underline{\mu}[\sigma(\underline{\mu})] \to \hat{e}(\sigma) \prec \Delta]$; so, $\hat{e}$ satisfies $\hat{K}_0 3$.

Thus, the $\hat{e}$ of our second form is also in $\hat{K}_0$.

Hence, by $\hat{K}3$, we have that $\hat{K} \subseteq \hat{K}_0$.

However, we can define the following function:

$$\delta(\sigma_\epsilon, \sigma) = \begin{cases} 1 \; if \;\; \sigma_\epsilon \equiv SE(0, 0, x) \\ \nabla \; otherwise \end{cases}$$

$$\delta(\sigma \sqcap \sigma', \sigma''') = Sup(\delta(\sigma, \sigma'''), \delta(\sigma', \sigma'''))$$

$$\delta(\sigma \sqcup \sigma', \sigma''') = Inf(\delta(\sigma, \sigma'''), \delta(\sigma', \sigma'''))$$

We then define $\hat{e} \equiv \lambda T.(\lambda x.\lambda T'.x) \downarrow_{T\delta(T,T)}$. This function, when applied to a knowledge state, checks if the first element is given to us in the form of an SE atom in the knowledge state; if so, then it outputs 1, otherwise, it will always output ∇. This function clearly satisfies the conditions to be in $\hat{K}_0$, but it is not of the form of a function available to $\hat{K}$ as, the δ defined above is not captured by our η-function, and hence $\hat{K} \neq \hat{K}_0$.

Thus $\hat{K} \subset \hat{K}_0$ as required. ♠

§6.5. Conclusion of Analysis

This section aims to summarise what we have achieved in this chapter. In essence, we began with our super-system and imposed the following restrictions.

1. Restricted our universe of choice sequences to satisfy OMR.

2. Further restricted our universe of choice sequences to satisfy s_{tri}.

3. The schema $\exists KS$-LS was imposed.

4. The schema $\hat{K}_0 1$ was weakened to $\hat{K}_0 1^*$.

5. We forced all predicates to be extensional (and assumed that extensionality was sufficient to remove additional choice sequences).

When these restrictions were imposed, our system reduced down to the conventional theory in §6.2 and §6.3, thus giving us a pathway to analysis.

§6.4 explored the idea of a series of (extensional) inductively defined functions that satisfy $\hat{K}_0$, the evaluation of which is computable. It also provided a way of evaluating functions that avoided the loop and contradiction cases explored at the end of §4.4.

7. Concluding Remarks

§7.1. Final Thoughts and Further Notions

This final chapter aims to summarise what we have so far achieved (§7.2) and what remains to be achieved (§7.3). §7.3 will be broken down further into specific aspects and questions that remain unanswered; any work deemed too incomplete to include in the main text above will also be explored here.

§7.2. The Current State of the Project

When we pause to reflect what we have achieved with this work, we find ourselves wondering if we have achieved anything at all. The existing theory offers us a clear path to analysis and certainly clarifies important technical points. However, we have covered some ground as seen below.

1 We have offered a compelling argument that the existing notions of lawlessness are insufficient (§3.7) and we have devised a new notion of lawlessness (§4.6) that is not only closed under continuous operations, but also evades Fletcher's Paradoxes (§4.9).

2 We have defined a language capable of expressing intensional information about choice sequences in a formal way (§4.3, §4.4 and §4.5). This language is readily extendible (see §7.3.2 for more on this) and it is also machine parsable (as seen in §6.4).

3 We have provided a deeper insight into the continuity schema, our axioms of knowledge proving to be the more primitive notions. These have allowed us to demonstrate just how far one may go with such basic ideas (§4.2.2).

4 The stronger of our two generalised stronger continuity schema (BC-N-KS^+ still proves the schema AC-NN, though we must omit the additional sequences for this proof to work (§4.8).

5 We have defined a more general notion of continuity (Σ-continuity) and a universe of functions ($\hat{K}_0$) capable of representing such Σ-continuous operations (§4.7).

6 We have constructed a generalised foundational system whose consistency, while unproven, is definitely plausible (§5.1); this system is both closed under continuous operations and contains a strong continuity principle as required by Dummett (1977).

7 This formal system was then exposed to a very clear set of restrictions and shown to reduce to an extension of $FIM\text{-}AN$. This has shed light on just what concessions we make to obtain analysis (§6.1, §6.2 and §6.3).

These, I feel, resolve the two issues we set out to resolve in the conventional literature; namely the confusion notions of lawlessness (1 in §3.9) and the lack of any real exploration of an intensional system (4 in §3.9).

These small refinements pave the way to a clearer understanding of the foundations of intuitionistic analysis. They illuminate areas of the theory that have previously been taken for granted, further reinforce the validity of existing work, and hint at a possible alternative intensional path to analysis. In essence, we have further cemented what has come before, and reached a little further on our stronger foundation.

§7.3. Further Refinements

This section aims to provide seven areas in which this work requires refinement; this text only represents the first step on what the author hopes will be a long and fruitful journey. §7.3.1 explores a possible strengthening of our axioms of knowledge; §7.3.2 looks at additional forms of knowledge, including the plausibility of higher order knowledge states; §7.3.3 looks closely at possible ways to clearly define a conception of contradictory knowledge; §7.3.4 looks at a possible extension to our treatment of spreads and bar induction; §7.3.5 looks at the assumptions we have made for analysis in our work; §7.3.6 highlights the absence of a consistency proof; §7.3.7 considers possible strengthenings of $\hat{K}$ and, §7.3.8 offers our closing remark.

§7.3.1. Axioms of Knowledge

In §4.2.2, we gave the two axioms of knowledge below.

σ-μ-1 $\forall \underline{\mu}[A(\underline{\mu}) \rightarrow \exists \sigma[\exists \underline{\nu}[\sigma(\underline{\mu}, \underline{\nu})] \wedge A'(\sigma)]]$, where A has no other free choice sequence variables, and $|\underline{\mu}|$ need not equal $|\underline{\nu}|$.

σ-μ-2 $\forall \sigma[A'(\sigma) \rightarrow \forall \underline{\mu}[\exists \underline{\nu}[\sigma(\underline{\mu}, \underline{\nu})] \rightarrow A(\underline{\mu})]]$, where A has no other free choice sequence variables, and $|\underline{\mu}|$ need not equal $|\underline{\nu}|$.

From these, we derived both a generalised form of open data (GOD) and a generalised form of weak continuity (WC-N-KS); however, we did not provide any such proof for our stronger continuity axiom (BC-N-KS) used in the system FIM-KS . This is because, at the moment, our axioms of knowledge are simply too weak to obtain the result we desire. The crucial line in the proof is the application of σ-μ-1 as shown below.

'By σ-μ-1 we have that $\exists \sigma[\exists \underline{\nu}[\sigma(\underline{\mu}, \underline{\nu})] \wedge A'(\sigma)]$' (Proof of GOD)

'By σ-μ-1 this implies that $\exists x \exists \sigma[\exists \underline{\nu}[\sigma(\underline{\mu}, \underline{\nu})] \wedge A'(\sigma, x)]$' (Proof of WC-N-KS)

The second of these two applications indicates that perhaps a strengthening of σ-μ-1 to something like,

σ-μ-1^+ $\forall \underline{\mu}[\exists x[A(\underline{\mu}, x)]] \rightarrow \exists \hat{e} \in \hat{K}_0 \forall \underline{\mu} \exists \sigma[\exists \underline{\nu}[\sigma(\underline{\mu}, \underline{\nu})] \wedge A'(\sigma, \hat{e}(\sigma))]]$, where A has no other free choice sequence variables, and $|\underline{\mu}|$ need not equal $|\underline{\nu}|$,

would provide us with a similar proof for BC-N-KS. The verbal translation of this is as follows.

'If given any $\underline{\mu}$ we can construct an x such we have a proof of $A(\underline{\mu}, x)$ then this implies that we can construct some extended neighbourhood function $\hat{e}$ such that, for any $\underline{\mu}$, we can construct a σ consistent with $\underline{\mu}$ (and possibly some other choice sequences $\underline{\nu}$) such that the knowledge is sufficient to prove A and also sufficient to construct our x'

This translation definitely shows that we have made a stronger statement here and it, implies a lot of power to our $\hat{K}_0$ function. The converse of this also seems very reasonable and we can definitely prove WC-N-KS and GOD with this schema; however, I am loath to adopt this knowledge schema in my theory without further exploration of its implications, especially since I am still doubtful of the potency of $\hat{K}_0$ in its current state. This is the area of work for a short paper, no doubt.

§7.3.2. Additional Forms of Knowledge

In §4.3 and §4.4, we introduced our atomic forms of knowledge, the formal restrictions we have in place to denote what information we will be considering. While our conception is certainly wider than in the conventional theory (indeed this is one of our main achievements); it is still an open question if additional forms of first order knowledge are required to give a full picture of the universe of choice sequences.

There are two additional notions that bear further exploration; the first is that we actually make tacit use of with the schema OMR, namely, the idea of higher order knowledge states. Is it possible to construct additional layers to the language of knowledge states so that we may faithfully represent these restrictions on knowledge states, and hence, devise a wider universe of choice sequences based on these second order restrictions? It certainly seems plausible, though a great many things do before one attempt to formalise the idea.

The second of these additional notions is placing a domain on knowledge; for example, given a sequence, what if we want to 'choose' the first ten terms freely and then adopt some law relation to generate the next ten, alternating between ten free choices and ten law bound choices? One may be tempted to assume that a spread already does this; however, because of the strong intensional nature of knowledge states, the extensional notion of spreads would fail to capture all the information one may wish to preserve. Thus, our current language of knowledge states is insufficient for this; however, some form of domain restriction certainly seems reasonable; for example, $LR(R, 0, 1, 9 < x < 29)$ or $LR(R, 0, 1, x \ mod \ 20 > 10)$. This notion certainly seems harmless enough, but again, I would hesitate to include it without much further exploration.

Overall, the treatment of atomic kinds of knowledge in this work is sufficiently strong to provide us with analysis, but quite insufficient to provide us with the absolute clarity on all choice sequences that we desire.

§7.3.3. Contradictory Knowledge States

Sadly, in our coverage of contradictory knowledge states, we are forced into a corner since the consistency of knowledge states is not decidable in general; however, we might obtain a better idea on just how decidable borderline cases are, if we had a better understanding of what it truly means for a knowledge state to be inconsistent.

We avoid having to define a solid notion of inconsistency for knowledge states by restricting our interest to Σ_{SE}, a species of knowledge state where inconsistency is decidable. A very informal idea is that 'a knowledge state is inconsistent iff it is impossible for it to be part of the generating process of a choice sequence (or tuple of)'. Formally, we would write this as follows.

$$\forall \sigma [\sigma \ is \ inconsistent \iff \neg \exists \underline{\mu} [\sigma(\underline{\mu})]]$$

This is clearly the right idea. However, it is also utterly useless to us in terms of identifying borderline cases. A slightly more useful notion would be the equivalent 'a knowledge is inconsistent iff its consistency with any given tuple of sequences would lead to absurdity'. Formally we would write this as follows.

$$\forall\sigma[\sigma \textit{ is inconsistent } \iff \forall\underline{\mu}[\sigma(\underline{\mu}) \to \bot]] \text{ (a)}$$

While this notion appears more interesting, it still does not give us a concrete way to identify inconsistent knowledge states. An overlooked tool may be our $\hat{K}_0$ functions, which would allow us to define (and make use of) something like the following.

$$\forall\sigma[\sigma \textit{ is inconsistent } \iff \exists\hat{e} \in \hat{K}_0[\hat{e}(\sigma) = \Delta]] \text{ (b)}$$

However, closer inspection shows us that this only works because we assert

$$\forall\sigma[\exists\underline{\mu}[\sigma(\underline{\mu})] \to \hat{e}(\sigma) \prec \Delta]$$

($\hat{K}_0 3$)

which makes us question if we are really just asserting the (weaker) statement below.

$$\forall\sigma[\sigma \textit{ is not inconsistent } \iff \exists\underline{\mu}[\sigma(\underline{\mu})]] \text{ (c)}$$

Another tool we could make use of is our $\hat{K}$-functions, essentially giving us the statement below.

$$\forall\sigma[\sigma \textit{ is inconsistent } \iff \exists\hat{e} \in \hat{K}[\hat{e}(\sigma) = \Delta]] \text{ (d)}$$

This is actually slightly stronger than (b), since it provides a (purely mechanical) method (as seen in §6.4) for identifying contradictory knowledge states; though here we hit a different kind of snag since our $\hat{K}$-functions are not capable of parsing certain kinds of knowledge at the moment (see more about this in §7.3.7), and thus, would render this to the weaker (though still useful) statement below.

$$\forall\sigma[\exists\hat{e} \in \hat{K}[\hat{e}(\sigma) = \Delta] \to \sigma \textit{ is inconsistent}]$$

Overall, I feel that our treatment of inconsistency in this work leaves much to be desired and is definitely an avenue for further work, though I believe its resolution may lie in providing a stronger notion of $\hat{K}$ that captures the full notion of $\hat{K}_0$ and not just its extensional component.

§7.3.4. Spreads and Bar Induction

Something the reader will have undoubtedly noticed as rather single minded is our treatment of spreads and bar induction. We deliberately adopt conventional notions to give ourselves an 'easy ride' to analysis. Specifically, we only work with spreads over Baire Space and not over the universe of knowledge states. Extending this notion of some form of 'knowledge spread' may offer us both a path to the higher order restrictions we desire; for example the case below.

$$S(\sigma) = \begin{cases} 0 \; if \; \sigma \in \Sigma_{SE} \\ 1 \; otherwise \end{cases}$$

It certainly matches the criteria to be a 'spread' (decidable, root admissible and infinite branches); and if we had some general method of constructing these spreads, then we could easily impose second order restrictions on choice sequences by restricting their knowledge states restricted to a knowledge spread.

This links directly to our second critique in this section, the treatment of bar induction. In §4.10 we introduced our schema for bar induction and specifically mentioned that we would only be considering extensional bars even after demonstrating the existence of a perfectly valid intensional bar: $R(\sigma) \iff \exists s[Spread(s) \in \sigma]$.

The main reason for this is that we were only looking at bar induction over Baire Space for the purpose of analysis; but what if we wished to establish a similar premise of our tree of knowledge? This notion would certainly be difficult to define due to our $\sqcup$ connective, though this is not to say that they would provide an insurmountable difficulty. A test schema I considered for the two specific bars mentioned previously is given below.

BI_D-KS^+

$$\{\forall\underline{\mu}\exists\sigma[\sigma(\underline{\mu}) \wedge R(\sigma)]$$

$$\wedge\, \forall\sigma[R(\sigma) \vee \neg R(\sigma)]$$

$$\wedge\, \forall\sigma[R(\sigma) \rightarrow A(\sigma)]$$

$$\wedge\, \forall\sigma[\exists x \exists w \forall y[A(\sigma \sqcap SE(w,x,y))] \wedge \forall s[A(\sigma \sqcap Spread(s))] \rightarrow A(\sigma)]$$

$$\rightarrow A(\sigma_{\emptyset})$$

However, I chose to leave it out of the main theory, as it only covers some very specific cases, of which there may well be more. The difficulty in attempting to further generalise the idea of line four of BI_D-KS^+ leads me to doubt the existence of a fully generalised notion of bar induction, though this does not rule out improving upon what we already possess.

Overall the treatment of spreads is intrinsically tied to the notion of bar induction and I hope that a resolution of one will perhaps shed some light upon the other. The grounds here are fertile for further exploration, especially should a full formalisation of second order knowledge states prove possible.

§**7.3.5. Analysis and Assumptions**

Our pathway to analysis (covered in §6.1, §6.2 and §6.3) was an exercise in following the path of least resistance; we made a great many assumptions which we reiterate below for the reader's convenience.

1. Restricted our universe of choice sequences to satisfy OMR.

2. Further restricted our universe of choice sequences to satisfy s_{tri}.

3. The schema $\exists KS$-LS was imposed.

4. The schema $\hat{K}_0 1$ was weakened to $\hat{K}_0 1^*$.

5. We forced all predicates to be extensional (and assumed that extensionality was sufficient to remove additional choice sequences).

Given the foundational nature of the subject this seems like a great many assumptions to impose, and, really only assumption (2) seems reasonable. Assumption (1) carries the same artificial flavour that we criticised KT-LL for, a strong argument for a path to analysis without it, in the author's point of view. Assumption (3), I believe, should be provable in a much more formal manner, but it relies on the schema OMR to obtain this proof. It is my opinion that, once again, a stronger understanding of second order knowledge states would shed some light upon this matter. Finally, my objections with assumptions 4 and 5 are linked to the objections the conventional literature raises against WC-N (see §2.4.1) and might conceivably be solved by replacing (5) by the imposition of graph extensionality (and thus (4) would not be needed). The existing works on graph extensionality are promising, as it has already been shown that graph extensionality is sufficient to prove pointwise continuity [Gielen, Swart and Veldman, 1981].

While our reduction to analysis is solid and sheds some light on the works in the existing literature, it is the author's feeling that too many concessions were made, and that the light we cast upon the matter was tainted by this. A full illumination of the matter will only be achieved when we are able to find a way to evade the need for the schema OMR (and hence $\exists KS$-LS) and reclaim UC under graph extensionality.

§**7.3.6. Consistency**

Something that is clearly lacking in the theory is the absence of any formal consistency proof. The reasons behind this are that attempting to use the function realisability method is likely to be extremely complex in this particular case and forming a Beth-Kripke model would also be challenging. The consistency of FIM-KS is certainly arguable; it merely generalises ideas expressed by existing consistent theories, and when subject to the restrictions imposed by analysis, it provides a very believable extension of FIM-AN (as seen in §6.3).

Each of the axioms in FIM-KS is vigorously argued for in §4; given their nature, I can see no reason why they shouldn't be consistent.

Exploring additional means of proving consistency as well as attempting those mentioned above is the first port of call for remedying this. The author is dubious if relative consistency is possible due to the strong intensional aspects of the theory.

§7.3.7. The Strength of $\hat{K}$

As mentioned in §7.3.3, the species of inductively defined extended neighbourhood functions is lacking in some ways; the forms of atomic knowledge the elements of $\hat{K}$ may parse are very limited, and they may only request elements rendering all such functions purely extensional.

The resolution to the first of these could be bought about by defining an η-function strong enough to extract elements from *Fan*, *Spread* and SR atoms. Should this be managed, we would have strong grounds to assume, for the extensional subset of $\hat{K}_0$ functions $\hat{K}_0E$, that $\hat{K} = \hat{K}_0E$. In essence, this would give us something just as powerful as a generalised induction over extensional bars.

The resolution of the second point requires us to identify all potential types of intensional knowledge requests, a far more difficult task. The success of this would give us grounds to argue, for the intensional subset of $\hat{K}_0$ functions $\hat{K}_0I$, that $\hat{K} = \hat{K}_0I$, and thus, give us something equivalent to a intensional bar induction.

Should we be fortunate to resolve both issues, then we would have good reason to assert $\hat{K} = \hat{K}_0$, and thus, something equivalent to our fully generalised bar induction over knowledge states. This may, additionally, pave the way to a proof of BC-N-$KS \vdash AC$-NN without needing the restrictions of analysis. This would, also, provide a solution to the issues raised in §7.3.4.

§7.3.8. A Final Thought on Interdependence

There is no doubt the reader has noticed that most of the future work expressed in this chapter is interlinked, and thus even the smallest of changes forces one to re-evaluate almost all of what one knows. It is this, I feel, that will prove the greatest challenge for anyone seeking to resolve the matters discussed above.

That the existing literature has blazed the path to analysis and marked it well is without doubt; but surely now, we have such a clear path it is time to explore some of the side roads and see if a shorter and more concise route may be found.

8. References

Berger U., Osswald H. and Schuster P. (2001): *Reuniting the antipodes - constructive and nonstandard views of the continuum.* Springer. Netherlands

Bernini S. (1976): A very strong intuitionistic theory. In:*Studia Logica* 35, vol 4, pp. 337–385

Bishop E. (1967): *Foundations of constructive analysis.* McGraw-Hill. USA

Bridges D. and Richman F. (1987): *Varieties of constructive mathematics.* Cambridge University Press. Cambridge

Brouwer L.E.J. (1907):*Over de grondslagen der wiskunde, Ph.D. thesis, University of Amsterdam.* English translation in Brouwer (1975) pp.11–101

Brouwer L.E.J. (1908):De onbetrouwbaarheid der logische principes. In:*Tijdschrift voor Wijsbegeerte* 2, pp.152-158. English translation in Brouwer (1975) pp.107–111

Brouwer L.E.J. (1912):Intuitionisme en formalisme. Inaugral address, University of Amsterdam. English translation by Arnold Dresden in Brouwer (1975) pp.123–138

Brouwer L.E.J. (1914): A. Schoenflies und H. Hahn. die entwickelung der mengenlehre und ihrer Anwendungen, Leipzig und Berlin 1913. In:*Jahresb. D.M.V.* 23, pp.78–83

Brouwer L.E.J. (1918): Begrundung der mengenlehre unabhngig vom logischen satz vom ausgeschlossenen dritten. zweiter teil, theorie der punktmengen. In:*KNAW Verhandelingen* 7. pp.1–33. English translation by Appleby J.F. and Rittberg C.J. (unpublished)

Brouwer, L.E.J. (1925): Zur berundung der intuitionistischen mathematij I. In:*Math Annalen* 93, pp.244–257

Brouwer, L.E.J. (1927): Uber Definitionsbereiche von Funktionen. In:*Mathematische Annalen* 97, pp.60–75. English translation of sections 1-3 in van Heijenoort (1967), pp.457–463

Brouwer, L.E.J. (1930): Die Struktur des Kontinuums, Wien: Komitee zur Veranstaltung von Gastvortrgen auslandischer Gelehrter der exakten Wissenschaften. English translation in Mancosu (1998), pp. 54–63

Brouwer, L.E.J. (1948A): Essentieel negatieve eigenschappen, In:*Indagationes Mathematicae* 10, pp 322–323. English translation in Brouwer (1975), pp478–479

Brouwer L.E.J. (1975): *Collected works (volume 1), A. Heyting (ed.).* Elsevier. Amsterdam/New York/Oxford

Brouwer L.E.J. (1981): *Brouwer's Cambridge Lectures on Intuitionism, D. van Dalen (ed.).* Cambridge University Press. Cambridge

Troelstra A.S. and Dalen D. (1982): *The L.E.J. Brouwer centenary symposium.* North Holland Publishing Company. Amsterdam

Troelstra A.S. and Dalen D. (1988): *Constructivism in mathematics; an introduction (2 volumes).* Elsevier. Amsterdam/New York/Oxford

Draglin A.G. (1988): *Mathematical Intuitionism: Introduction to Proof Theory.* American Mathematical Society

Dummett M.A.E. (1977): *Elements of intuitionism.* Clarendon Press. Oxford

Fletcher P. (1998): *Truth proof and infinity ; a theory of constructions and constructive reasoning.* Kluwer. Netherlands

Heyting A. (1930): *Die formalen regeln der intuitionistischen logik: die formalen regeln der intuitionistischen mathematik.* Akademie der Wissenschaften [und] De Gruyter

Heyting A. (1934): it Mathematische grundlagenforschung. Intuitionismus. Bewisthoeorie. Springer Verlag. Berlin

Horsten L. and Welch P. (2016): *Godel's disjunction: the scope and limits of mathematical knowledge.* Oxford Press

Gielen W. De Swart H. and Veldman W. (1981): The continuum hypothesis in intuitionism. In*Journal of Symbolic Logic* 46, pp.121–135

Goodman N.D. (1979): Review of Dummett (1977). In:*Journal of Symbolic Logic* 44, pp.276–277

Kachapova F. (2015): A strong multi-typed intuitionistic theory of functionals. In*Journal of Symbolic Logic* 80, pp.1035–1065

Kleene S.C. and Vesley R.E. (1965): *The foundations of intuitionistic mathematics, especially in relation to recursive functions.* North Holland. Amsterdam

Kolmogorov A.N. (1932): Zur deutung der intuitionistischen logic. In:*Mathematische Zeitschrift* 35, pp.58–65

Kreisel G. (1958): A remark on free choice sequences and the topological completeness proofs. In:Journal of Symbolic Logic 23, pp.369–388

Kreisel G. (1967): Informal rigour and completeness-proofs. In:*Lakatos 1967* pp.138–186

Kreisel G. (1968): Lawless sequences of natural numbers. In: *Compositio Math* 20, pp. 222–248

Lakatos I. (1967): *Problems in the Philosophy of Mathematics: Proceedings of the International Colloquium in the Philosophy of Science, London, 1965.* Elsevier

Martino E. (1982): Creative Subject and Bar Theorem. In:*Troelstra and Dalen (1982)* pp.311–318

Mancosu P. (1998): *From Brouwer to Hilbert ; The debate on the foundations of mathematics in the 1920s.* Oxford University Press

Moschovakis J.R. (1987): Relative Lawlessness in Intuitionistic Analysis. In:*Journal of Symbolic Logic* 52. pp.68–88

Moschovakis J.R. (1993): An intuitionistic theory of lawlike, choice and lawless sequences. In:*Logic Colloquium '90: ASL Summer Meeting in Helsinki*, pp.191–209

Moschovakis J.R. (1994): More about relatively lawless sequences. In:*Journal of Symbolic Logic* 59. pp.813–829

Moschovakis J.R. (2016): Iterated definability, lawless sequences and Brouwer's continuum. In:*Horsten and Welch (2016)*

Myhill J. (1966): Notes towards an axiomatization of intuitionistic analysis. In*Logique et Analyse(N.S.)* 9, pp.280–297

Myhill J. (1968): Formal systems of intuitionistic analysis I. In:*Rootselaar and Stall (1968)* pp.161–178

Niekus J. (2005): Individual choice sequences in the work of L.E.J Brouwer. In*Philosophia Scientiae cahier special* 5, pp.217–232

Niekus J. (2010): Brouwer's incomplete objects. In*Hitsory and Philosophy of Logic* 31, pp.31–46

Niekus J. (2017): What is a choice sequence? How a solution of Troelstra's paradox shows the way to an answer to this question. (Unpublished)

Rootselaar B. and Staal J. (1968): *Logic, Methodology and Philosophy of Science III.* North-Holland Publishing Company. Amsterdam

Troelstra A.S. (1968): The theory of choice sequences. In: *Rootselaar and Staal (1968)* pp.201–223

Troelstra A.S. (1969): Principles of intuitionism. In: *Lecture Notes in Mathematics* 95. Springer-Verlag. Berlin

Troelstra A.S. (1969A): Informal theory of choice sequences. In: *Studia Logica* 25, pp31–54

Troelstra A.S. (1977): *Choice sequences : a chapter of intuitionistic mathematics.* Clarendon Press. Oxford

Troelstra A.S. (1982): On the origin and development of Brouwer's concept of choice sequence. In:*Troelstra and Dalen (1982)* pp465–486

Veldman W. (2001): Understanding an using Brouwer's continuity principle. In:*Berger, Osswald and Schuster (2001)* pp.285-302

Veldman W. (2006): Brouwer's real thesis on bars. In:*Philosophia Scientiae* 6, pp.21–42

Veldman W. (2008):Some applications of Brouwers thesis on bars. In:*Atten, Boldini, Bourdeau and Heinzmann (2008)* pp.326–340

Van Atten M. and Van Dalen D. (2002): : Arguments for the continuity principal. In*Bulletin of Symbolic Logic* 8, pp.329–347

Van Atten M. Boldini P. Bourdeau M. and Heinzmann G. (2008): *One hundred years of intuitionism (1907-2007).* Birkhauser. Basel/Boston/Berlin

Van Atten M. (2010): *Brouwer meets Husserl:on the phenomenology of choice sequences.* Springer. Dordrecht

Van Dalen D. (1968): Intuitionsitic Logic. In*The handbook of Philosophical Logic* 66, pp.225–339

Van Dalen D. (1999): *Mystic, geometer, and intuitionist: the life of L.E.J Brouwer (volume 1).* Clarendon Press. Oxford

Van Heijenoort J. (1967):*From Frege to Godel. a sourcebook in mathematical logic, 1879–1931*, Cambridge (MA): Harvard University Press

Van Der Hoeven. (1982): *Projections of lawless sequences.* Mathematisch Centrum

Van Stigt W. (1990):*Brouwer's intuitionism.* North Holland Publishing Company. Amsterdam

9. Index of Words and Phrases

10. Index of Symbols (Alphabetical)

11. Index of Symbols (Chronological)

12. Index of Schemata and Axioms

13. Index of Theorems and Lemmas

14. Index of Examples

15. Index of Formal Theories